11-2023

Calorimetric Analysis Using DNA Thermal Stability to determine protein concentration

Matthew W. Eskew
Portland State University, meskew@pdx.edu

Patrick Reardon
Oregon State University

Albert S. Benight
Portland State University, abenight@pdx.edu

Citation Details

Publshed as: Eskew, M. W., Reardon, P. W., & Benight, A. S. (2023). Calorimetric analysis using DNA thermal stability to determine protein concentration. bioRxiv, 2023-09.

Calorimetric analysis using DNA thermal stability to determine protein concentration

Authors: Matthew W. Eskew[1,2*], Patrick W. Reardon[3] and Albert S. Benight[1,2,4]

Affiliations:

[1]ThermoCap Laboratories Inc, Portland, Oregon
[2]Department of Chemistry, Portland State University, Portland, Oregon
[3]OSU NMR Facility, Oregon State University, Corvallis, Oregon
[4]Department of Physics, Portland State University, Portland, Oregon

* Corresponding Author meskew@thermocaplabs.com

ABSTRACT

It was recently reported for two globular proteins and a short DNA hairpin in NaCl buffer that values of the transition heat capacities, $C_{p,DNA}$ and $C_{p,PRO}$, for equal concentrations (mg/mL) of DNA and proteins, are essentially equivalent (differ by less than 1%). Additional evidence for this equivalence is presented that reveals this phenomenon does not depend on DNA sequence, buffer salt, or T_m. Sequences of two DNA hairpins were designed to confer a near 20°C difference in their T_m's. For the molecules, in NaCl and CsCl buffer the evaluated $C_{p,PRO}$ and $C_{p,DNA}$ were equivalent. Based on the equivalence of transition heat capacities, a calorimetric method was devised to determine protein concentrations in pure and complex solutions. The scheme uses direct comparisons between the thermodynamic stability of a short DNA hairpin standard of known concentration, and thermodynamic stability of protein solutions of unknown concentrations. In all cases, evaluated protein concentrations determined from the DNA standard curve agreed with the UV-Vis concentration for monomeric proteins. For samples of multimeric proteins, streptavidin (tetramer), Herpes Simplex Virus glycoprotein D (trimer/dimer), and a 16 base pair DNA duplex (dimer), evaluated concentrations were greater than determined by UV-Vis by factors of 3.94, 2.65, and 2.15, respectively.

INTRODUCTION

Classical techniques for determining concentrations of DNA and protein solutions include absorbance spectroscopy (UV-Vis), the BCA, Bradford, Lowry and fluorescence assays which have the added requirement of specific molecular probes [1-3]. UV-Vis, either from direct measurements or a colorimetric assay, is the most widely employed method for quantitative determination of concentrations of nucleic acids and proteins in solution. To determine accurate protein concentrations for a specific molecule, these methods typically require pure solutions. Even though standard methods can generally provide reliable measurements of protein concentrations, they can yield conflicting results [4]. For many cases UV-Vis measurements can provide accurate concentrations. However, for complex solutions, such as those encountered in cell culture media and lysates, reliable concentrations from UV-Vis measurements can be tenuous and generally only provide a semi-quantitative estimate of the total protein concentration.

It has been well established that thermodynamic analysis of the temperature induced melting transitions of proteins and DNA provides a novel means to explore many aspects of molecular structural composition [5-13]. Calorimetric measurements of the transition provide a direct reflection of the collective forces responsible for molecular stability [6, 11, 14-16]. Many of the features related to molecular processing in biology such as ligand binding and intermolecular interactions can be probed, with unique information extracted, using calorimetric measurements [16-18].

Thermal denaturation or melting experiments are performed in a differential scanning calorimeter (DSC). In DSC melting experiments for DNA or proteins the heat capacity, C_p, at constant pressure is measured as a function of temperature. A plot of C_p versus T is the DSC melting curve or thermogram. For DNA and many proteins, the heat induced melting transitions are endothermic with a pseudo-Gaussian shape. The temperature at the maximum peak height is defined as the transition temperature, T_m. Likewise, the value of C_p at the maximum peak height is the transition heat capacity.

Here we report on the use of melting calorimetry to determine unknown concentrations of protein solutions. It is founded on the recent discovery for equal mass concentrations (mg/mL) of DNA or protein that their heat capacities, determined from the DSC thermogram peak heights at the transition temperatures, are the same i.e. $C_{p,PRO} = C_{p,DNA}$ [19]. Based on this relationship a novel scheme was developed that enabled quantitatively accurate determinations of targeted protein concentrations in both pure solutions and impure complex mixtures.

Readily available synthetic short DNA oligomers with virtually any sequence provide the opportunity to fine-tune, by design, DNA secondary structure and corresponding sequence dependent thermodynamic stability. For example, several sequence dependent features of the thermodynamic stability of short DNA intramolecular hairpins have been well characterized [7,

20, 21]. As shown here, a DNA hairpin was chosen to serve as the stability standard for comparisons.

For the specifically designed and well characterized intramolecular hairpin that was used, DSC thermograms were measured for the DNA as a function of concentration over a 100-fold range. Plotting the thermogram peak heights versus DNA concentration produced a standard curve. By comparing the peak height on the thermogram for a protein at unknown concentration to the standard curve, and accounting for differences between the partial specific volumes of proteins and DNA, it is possible to directly evaluate the concentration of the protein.

MATERIALS AND METHODS

Reagents: Buffer solutions contained either NaCl or CsCl at 150 mM in 10 mM potassium phosphate, 15 mM sodium citrate adjusted to pH = 7.4 with HCl. Pure Proteins: Human Serum Albumin (HSA) (≥ 99% pure, Lot number: SLBT8667), Lysozyme (recombinant, expressed in rice, Lot number: SLCH2681), and Plasma (Human, Lot number: SLBT0202) were purchased from Sigma Aldrich (St. Louis, MO) and received as lyophilized power. Lysate from *E. coli* was purchased from Bio-Rad (Hercules, CA) and received as a lyophilized powder. The above proteins were prepared in buffer and stored at 4°C for at least 24 hours before use. Protein concentrations were confirmed by UV-Vis at A_{280} [22, 23].

Samples purchased from RayBiotech (Peachtree Corners, GA) included: HEK293 cellular supernatant, Lot number: 03U27020D. Unpurified angiotensin converting enzyme 2 (ACE2) expressed in human embryonic kidney cells (HEK293) received in the preparation media Lot number: 02U2802LL. Isolated, purified ACE2 Lot number: 04U27020GC. Purified receptor binding domain (RBD) from SARS-CoV-2, Lot Number: 05U22020TWB. Purified Human Herpes Simplex Virus glycoprotein D (HSV-GpD), Lot Number: 01U25022AG. All HEK293 solutions had a volume of 200 μL.

DNA: Two DNA hairpins were designed to have T_m's that differed by approximately 20 °C. The "high-temp" hairpin was formed from the 20-base sequence 5'-CGG GCG CGT TTT CGC GCC CG-3'. The "low-temp" hairpin also had 20-bases with the sequence 5'-CGA TCG CGT TTT CGC GAT CG-3'. DNA strands were purchased from IDT (Coralville, IA) and received following their standard desalting routine. Lyophilized DNA was resuspended in either buffered NaCl or CsCl solutions and stored at 4°C. DNA concentrations were determined spectrophotometrically at A_{260} and agreed with vendor specifications. Figurative structures and sequences for the high- and low-temp hairpins are depicted in Figure 1(a, b).

The 16 base-pair duplex DNA was formed from two individual stands whose base sequences are complementary. Duplex strands were purchased from IDT (Coralville, IA). The 5' sequence of the duplex is 5'-CGA TCA TCG ATG ATG C-3'. Following their standard desalting routine, the preformed duplex was received as a lyophilized powder. The double-stranded DNA was

resuspended in the desired appropriate buffer and stored at 4°C. DNA concentration was determined by UV-Vis at A_{260}. The duplex is depicted in Figure 1(c).
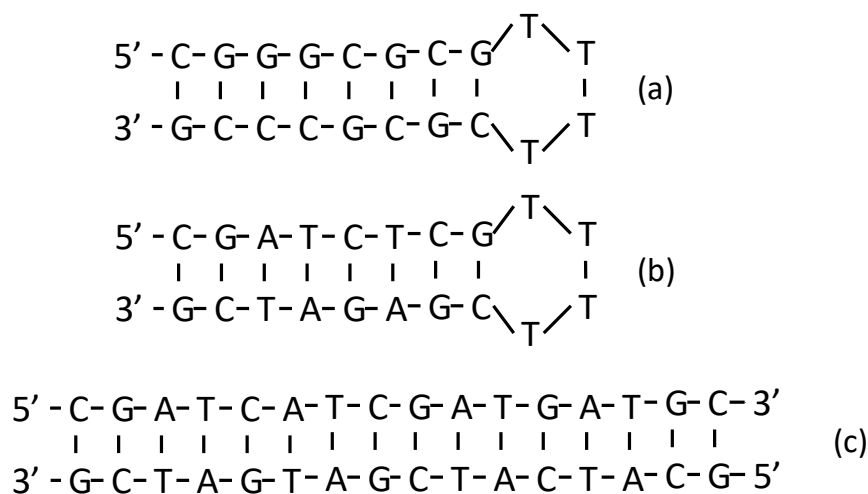

5' - C- G- G- G- C- G- C- G⟨T⟩T
    | | | | | | | | | (a)
3' - G- C- C- C- G- C- G- C⟨T⟩T

5' - C- G- A- T- C- T- C- G⟨T⟩T
    | | | | | | | | | (b)
3' - G- C- T- A- G- A- G- C⟨T⟩T

5' - C- G- A- T- C- A- T- C- G- A- T- G- A- T- G- C- 3'
  | | | | | | | | | | | | | | | | (c)
3' - G- C- T- A- G- T- A- G- C- T- A- C- T- A- C- G- 5'

Figure 1: DNA Structures: (a) high-temp hairpin, (b) low-temp hairpin, and (c) duplex DNA

Calorimetric Measurements: Melting experiments were performed on a CSC differential scanning microcalorimeter (now T.A. instruments, New Castle, DE). Samples were prepared by adding specific components to buffer or pre-prepared media solutions in a sample volume of 0.5 mL. For DSC melting experiments, the sample heating rate was approximately 1 °C/min while monitoring changes in the excess heat (microcalories) of the sample versus temperature [16, 24, 25]. In their primary form, melting curves or thermograms of protein and DNA samples were displayed as plots of changes in microcalories versus temperature. For complex media experiments all DSC measurements were made on samples residing in the lysate media, either *E. coli* lysate or HEK923 supernatant. DSC thermograms were measured as a function of concentration for the high and low-temp hairpins, HSA and lysozyme. For each experiment, the transition temperature, T_m, was determined by the temperature of the maximum peak height, $C_p^{max'}$ on the thermogram.

Data Reduction and Analysis: In primary form thermograms of proteins are displayed as plots of microcalories (μcal) versus temperature. Baselines of all thermograms were treated in an identical manner. Initially, for all mixtures the buffer baseline was first subtracted from the raw curves. Then baselines of the raw thermograms of ΔC_P or μcal versus temperature were determined using a four-point polynomial fit, over the temperature range of the transition. The result of the analysis produced baseline corrected thermograms used for comparisons and analysis [16, 17, 19]. For all thermograms of proteins in complex media, the thermogram for media alone served as the background that was then subtracted from thermograms of samples containing the target protein in media [18]. These difference thermograms were then used to evaluate the concentration of desired protein components in the mixture.

Statistical Analyses: For nearly all samples thermograms were collected at least in duplicate. Reproducibility and accuracy were verified using an ANOVA p-test. Statistical analysis and curve fitting were performed using Origin (Pro) Software, Version 2021b (Origin Corporation, Northhampton, MA). Linear fits of plots were quantified using a Pearson test [26, 27].

Partial Specific Volumes of DNA. PSV values for the DNA hairpins and duplex DNA were determined by Analytical Ultracentrifugation (AUC). All sedimentation equilibrium experiments were performed in a Beckman Coulter Optima XL-A analytical ultracentrifuge (Beckman Coulter, Brea, CA) at 20˚C equipped with absorbance optics. Data were measured at 260 nm in 6-channel cells. Samples were centrifuged at 24,000, 32,000, and 40,000 rpm, with a minimum equilibration time of 30 hours per speed. Buoyant molecular weights were determined using Heteroanalysis and PSV values were calculated using the Svedberg equation [28]. All experiments were conducted in standard NaCl buffer, with a density of 1.005584 g/ml.

RESULTS

The Transition Heat Capacity, $C_{p,x}$. Under appropriate experimental conditions and standard data analysis procedures, the maximum peak height on measured thermograms of ΔC_P versus temperature, $C_p^{max'}$, for x = DNA or proteins (PRO) can be given by [19, 29],

$$C_p^{max'} = (C_{p,x} - PSV_x) \cdot \rho_x \qquad (1)$$

Where ρ_x is the mass concentration, $C_{p,x}$ and PSV_x are the *transition heat capacity* and partial specific volume, respectively. It should be noted that ρ_x is the mass concentration of the sample reported as mg/mL. For the DNA and protein samples examined, an identical procedure was followed to evaluate $C_{p,x}$. Eqn (1) indicates a linear plot of $C_p^{max'}$ versus concentration, ρ_x, has a slope of $C_{p,x} - PSV_x$ [19]. For these samples thermograms were measured as a function of DNA concentration over the approximately 100-fold range from 0.02 to 2.5 mg/mL. All molecules were melted in buffer containing NaCl. To evaluate whether buffer composition played a role in the equivalence of transition heat capacity, some were also melted in buffer with CsCl replacing NaCl. For every sample, plots of $C_p^{max'}$ versus ρ_x were constructed. From linear fits of these plots, $C_{p,x}$ values were determined. Samples that were analyzed are listed in Table 1 along with results of the linear fits. As seen in Table 1, R^2 values (Pearson test) were greater than 0.99 in all cases, indicative of highly accurate linear fits. Note, these fits were not forced to pass through the origin. This turned out to be of little consequence since values reported in Table 1 were statistically equivalent for fits forced through zero. These results are shown in Table S1 of the Supplemental Material.

From the slopes given in Table 1 values of $C_{p,x}$ were determined for the DNA and protein samples. These are listed in Table 2. Previously, similar results were reported for the high-temp hairpin, HSA and lysozyme [19]. For the present study, new data was collected for these molecules. Where repeat experiments were conducted (high-temp hairpin, HSA and lysozyme in NaCl buffer) results were identical to those previously reported [18].

5

Table 1: Linear Fits				
Sample	PSV (mL/g)	Slope (µcal mL/mg)	Intercept	R²
High Temp DNA Hairpin (NaCl)	0.557	2.310 ± 0.033	0.119 ± 0.044	0.9988
High Temp DNA Hairpin (CsCl)	0.557	2.337 ± 0.057	0.040 ± 0.096	0.9982
Low Temp DNA Hairpin (NaCl)	0.557	2.322 ± 0.005	0.070 ± 0.010	0.9999
Low Temp DNA Hairpin (CsCl)	0.557	2.348 ± 0.095	0.050 ± 0.173	0.9951
Duplex DNA (NaCl)	0.625	4.140 ± 0.024	-0.082 ± 0.041	0.9999
Human Serum Albumin (NaCl)	0.733	2.122 ± 0.032	-0.098 ± 0.033	0.9996
Human Serum Albumin (CsCl)	0.733	2.248 ± 0.098	-0.156 ± 0.129	0.9962
Lysozyme (NaCl)	0.703	2.115 ± 0.104	-0.256 ± 0.108	0.9951

As depicted in Fig 1, both 20 base DNA strands form intramolecular hairpins with an 8 base-pair stem with a T_4 loop. For this study, the DNA duplex region of the original high-temp hairpin was re-designed to be relatively less stable by substituting A-T base pairs for G-C base pairs in the original sequence. The new sequence conferred a much lower T_m = 75.72 °C for the low-temp hairpin, compared to 94.44 °C. Thermograms measured as a function of concentration for the high-temp hairpin are shown Fig 2a. Similar plots for the other samples are given in the Supplemental Material. Note, for each hairpin the T_m was identical at all concentrations, behavior that is consistent with melting of a unimolecular tertiary structure. As shown in Fig 2a, peak heights, $C_p^{max'}$, for the high-temp hairpin exhibited a linear response to DNA concentration. This is clear in Fig 2b where $C_p^{max'}$ values for the high-temp DNA hairpin are plotted versus concentration.
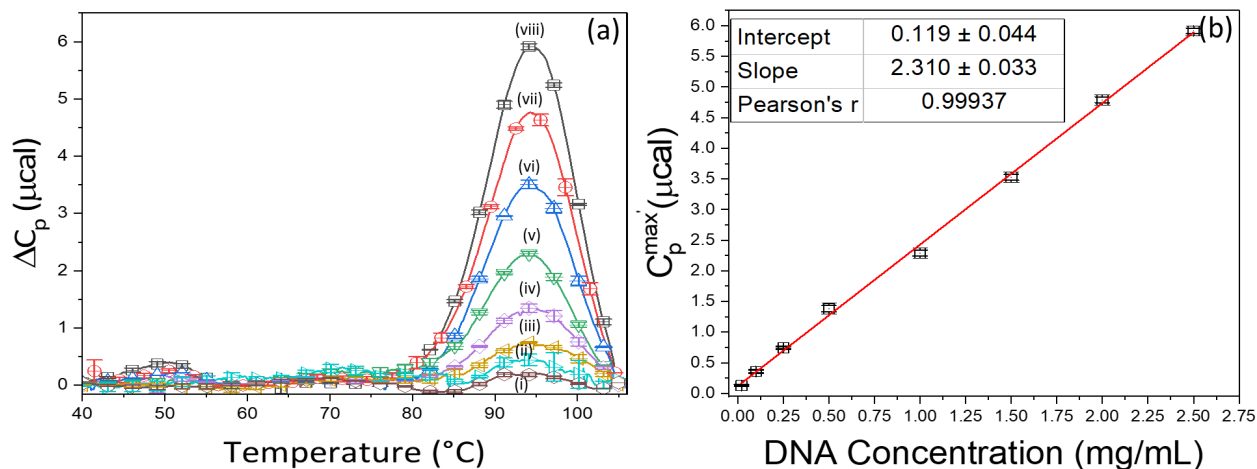

Figure 2: (a) Average thermograms for high-temp DNA in mg/mL: (i) 0.02, (ii) 0.10, (iii) 0.25, (iv) 0.50, (v) 1.00, (vi) 1.50, (vii) 2.00, (viii) 2.50. (b) Transition heat capacity curve for high-temp DNA.

From the slope of these linear plots (Table 1) and PSV for DNA it is possible to determine the transition heat capacity $C_{p,DNA}$. Reported PSV values vary for DNA from 0.57 ± 0.03 mL/g with the standard value assumed to be 0.55 mL/g [30, 31]. From AUC measurements for the DNA hairpins, PSV$_{DNA}$ = 0.557 ± 0.025 mL/g. With this value the transition heat capacity $C_{p,DNA}$ = 2.86 ± 0.07 mcal g^{-1} K^{-1}. Thermograms and linear plots like those shown Fig 2 were also obtained for the low-temp hairpin and HSA (Figs S1 and S2, Supplemental Material). The $C_{p,DNA}$ values for

6

these samples were in precise agreement with the value for the high-temp hairpin found here, and reported earlier for different samples of the same DNA and HSA [19]. Despite differences in their sequences, and corresponding different thermodynamic stabilities, $C_{p,DNA}$ values for the DNA hairpins were equivalent and independent of both sequence and T$_m$ and equal to $C_{p,PRO}$ for HSA at the same concentration [19].

NaCl versus CsCl: Effects of electrolyte size and/or mass on the values of $C_{p,DNA}$ and $C_{p,PRO}$ were investigated by performing analogous experiments with CsCl replacing NaCl in the melting buffer. For these experiments, stock solutions of the high- and low-temp DNA hairpins and HSA were also prepared in CsCl buffer. In the same manner as in NaCl buffer, thermograms were measured as a function of DNA concentration from 0.5 to 2.5 mg/mL. Then for the molecules in CsCl, plots of $C_p^{max'}$ versus concentration were constructed (as shown in Fig 2b for the high-temp hairpin in NaCl buffer).

Slopes of plots of $C_p^{max'}$ versus concentration yielded essentially identical $C_{p,DNA}$ values for the high-temp (2.88 ± 0.06 mcal g^{-1} K^{-1}) and low-temp (2.89 ± 0.09 mcal g^{-1} K^{-1}) hairpins in CsCl buffer. As displayed in Table 2, $C_{p,DNA}$ and $C_{p,PRO}$ values were in precise agreement with those measured for the DNA hairpins and for the HSA sample in CsCl, $C_{p,PRO} = 2.89 ± 0.09$ mcal g^{-1} K^{-1}. These values of $C_{p,DNA}$ and $C_{p,PRO}$ determined in buffered NaCl or CsCl were identical. It should be noted in Table 2 for Streptavidin, HSV-GpD, and duplex DNA, evaluated transition heat capacities, $C_{p,PRO}$ and $C_{p,DNA}$ were 9.19, 6.33 and 5.50 mcal·g^{-1}·K^{-1}, factors of 3.19, 2.2 and 1.91 larger, respectively, than the average over the other molecules. For streptavidin and HSV-GpD samples $C_p^{max'}$ values measured at a single concentration were used to generate a full curve of $C_p^{max'}$ versus concentration by linear interpolation through the origin. From the slope of these plots $C_{p,PRO}$, in Table 2, was evaluated. As posited below, the apparent discrepancies for streptavidin, HSV-GpD and duplex DNA were likely due to their multimeric composition.

Table 2: Transition heat capacity $C_{p,x}$ for DNA and proteins in standard (NaCl) and cesium chloride (CsCl) buffers

Molecule	T$_m$ (°C)	$C_{p,x}$ (mcal·g^{-1}·K^{-1})
High Temp DNA Hairpin (NaCl)	94.44 ± 0.09	2.86 ± 0.03
High Temp DNA Hairpin (CsCl)	94.36 ± 0.06	2.88 ± 0.06
Low Temp DNA Hairpin (NaCl)	75.72 ± 0.13	2.87 ± 0.01
Low Temp DNA Hairpin (CsCl)	75.56 ± 0.06	2.89 ± 0.09
Duplex DNA (NaCl)	71.55 ± 0.07	5.50 ± 0.05
Human Serum Albumin (NaCl)	63.85 ± 0.14	2.86 ± 0.03
Human Serum Albumin (CsCl)	63.16 ± 0.13	2.89 ± 0.09
Lysozyme (NaCl)	64.77 ± 0.08	2.82 ± 0.13
Streptavidin (NaCl)	77.39 ± 0.05	9.19
HSV – glycoprotein D (NaCl)	57.46 ± 0.02	6.33

Determination of Unknown Protein Concentrations: With $C_{p,DNA} = C_{p,PRO}$, a scheme was devised to determine unknown concentrations of proteins in both pure and complex media solutions. Central to this process is construction of the standard curve. In the present case, from the collection of plots of $C_p^{max'}$ versus concentration that were constructed for the DNA hairpins and HSA, the plot for the high-temp hairpin was chosen as the standard curve. This choice is somewhat arbitrary. Analogous plots for the other molecules could have served equally as well. Once the standard curve was constructed, the sole input required to determine the unknown concentration of a protein is the peak height from a single thermogram for the target protein. The unknown concentration of the protein is the point on the ordinate axis on the DNA standard curve corresponding to $C_p^{max'}$ of the protein on the abscissa of the standard curve.

Accurate evaluations of protein concentrations, by comparison of diagnostic results for protein solutions of unknown concentrations with the DNA standard curve, also requires a correction that accounts for differences between PSV values for DNA and proteins. While the transition heat capacities for HSA and lysozyme, $C_{p,PRO}$, and DNA, $C_{p,DNA}$ in Table 2 are essentially equivalent, the average PSV for proteins are different. As measured by AUC, the PSV$_{DNA}$ = 0.557 ± 0.025 mL/g while most globular proteins have PSV values that vary from 0.70-0.75 mL/mg [32-36]. Also, for independent lots of duplex DNA the evaluated PSV was 0.623 ± 0.004 mL/g, considerably larger than what has been reported for duplex DNA [30]. In fact, differences in the slopes of plots of $C_p^{max'}(protein)$ or $C_p^{max'}(DNA)$ versus concentration (Fig 2b, Table 1) are due directly to differences between PSV values for proteins and DNA.

After accounting for differences in PSV values for DNA and protein, the protein concentration is obtained. As will be shown below this process provides semi-quantitative estimates of protein concentrations in both pure and complex media.

From rearrangement of Eqn (1) the *transition heat capacity* is given by,

$$C_{p,x} = \frac{C_p^{max'}}{\rho_x} + PSV_x \qquad (2)$$

The analytical process is founded on the equality of values of the transition heat capacities for protein and DNA, determined by Eqn (2). Since $C_{p,PRO} = C_{p,DNA}$,

$$\frac{C_{p,PRO}^{max'}}{\rho_{PRO}} + PSV_{PRO} = \frac{C_{p,DNA}^{max'}}{\rho_{DNA}} + PSV_{DNA} \qquad (3)$$

Solving for the unknown protein concentration,

$$\rho_{PRO} = \frac{C_{p,PRO}^{max'}}{\frac{C_{p,DNA}^{max'}}{\rho_{DNA}} + (PSV_{DNA} - PSV_{PRO})} \qquad (4)$$

8

Eqn (4) reveals protein concentration, ρ_{PRO}, can be evaluated from a single $C_{p,PRO}^{max'}$ for a target protein. Practically, reasonable estimates of protein concentration can be obtained from a single DNA thermogram of known concentration. Provided that $C_{p,DNA}^{max'}$ is taken from DNA thermograms measured at concentrations above 1.0 mg/mL since larger deviations occur at smaller DNA concentrations.

More accurate concentration determinations can be made using the complete DNA standard curve. The standard curve displayed in Fig 2b was constructed from multiple measurements of thermograms for the DNA standard as a function of concentration. The slope of the best fit line to the standard curve, $C_{p,DNA}^{max'}/\rho_{DNA}$, was evaluated and inserted in Eqn (4). Then after accounting for the difference between the partial specific volumes for the DNA and protein, $PSV_{DNA} - PSV_{PRO}$, concentration of the protein ρ_{PRO}, was determined from $C_{p,PRO}^{max'}$.

Eqn (4) was employed to determine the protein concentrations in several solution contexts. In what follows, examples include proteins in pure solutions, plasma and complex fluids such as cellular media. Sample preparation and experimental and analytical protocols were as described in Materials and Methods.

Pure Solutions: The DNA standard curve displayed in Fig 2b was used to determine mass concentrations for several proteins and a few complex fluids. Some of which were also examined previously [18]. PSV values given in Table 3 for HSA [16, 19, 22], lysozyme [2,4], SARS-CoV-2 RBD [37-39], ACE2 [37], streptavidin [40], and HSV-GpD [41] were taken from the literature. Table 3 shows for all cases examined, there was excellent agreement between protein concentrations estimated from the DNA standard curve (column 3) and concentrations determined by UV absorbance or measured dry weight (column 4). Instead of an average over all peak heights at a particular concentration, peak heights in Table 3 were randomly selected from a set of peak heights obtained from the ranges of concentrations measured for each sample.

Table 3: Protein Concentration Comparison				
Sample	Peak Height (μcal)	PSV (mL/g)	DNA Assay (mg/mL)	Independent Concentration (mg/mL)
HSA	3.70	0.733	1.74	1.79
Lysozyme	3.55	0.703	1.65	1.80
SARS-CoV-2 RBD	0.75	0.71	0.35	0.40
ACE2	0.49	0.71	0.23	0.20
HSA in Plasma	2.08	0.733	0.98	1.08
E. coli Lysate	3.89	0.725	1.82	2.00
ACE2 unpurified	(C)0.09, (N) 0.06	0.69	(C) 0.04, (N) 0.03	------
Streptavidin	6.71	0.70	3.11 (3.94x)	0.79
HSV-GPD	2.23	0.74	1.06 (2.65x)	0.40
16mer dsDNA	4.09	0.623	1.83 (2.15x)	0.85

Overall, resulting concentrations evaluated from DSC measurements of proteins and the DNA standard curve differed (on average) by less than 8% from those determined by UV-Vis or dry weight.

In a prior study, two additional pure proteins, SARS-CoV-2 receptor binding domain (RBD) and angiotensin converting enzyme 2 (ACE2) were also examined [37-39]. Evaluated concentrations for these proteins agreed with independently determined reported values. Examination of Table 3 reveals for most samples, our method slightly underestimated the protein concentrations. The exception is ACE2 where the concentration was slightly overestimated by 13%. As previously noted, purified ACE2 samples displayed a single peak on their thermograms, in contrast to the expected two-peak behavior exhibited by native ACE2 [42]. Regardless, Table 3 indicates good agreement for concentration determinations for a variety of proteins in pure and complex media solutions, and DNA. Exceptions were observed for Streptavidin, HSV-GpD and short duplex DNA. As suggested below these exceptions are largely attributed to their higher molecular stoichiometries compared to the other molecules in Table 3.

Streptavidin is known to be a homo-tetramer comprised of four identical subunits (stoichiometry = 4) associated in a stable complex at 20°C regardless of concentration [43]. In Table 3 the concentration evaluated for streptavidin using the DNA assay (3.11 mg/mL) was also nearly a factor of four times (3.94) greater than the UV-Vis concentration of 0.79 mg/mL. To independently verify the streptavidin concentration, the commonly accepted extinction coefficient of 139,000 $M^{-1}cm^{-1}$ at 280nm was employed [43-45]. This value was used to determine the independent concentration of streptavidin given in column 4 of Table 3, also in agreement with the dry-weight mass used to prepare the protein solution. This observation is at odds with the calorimetric method (DNA assay) and apparently suggests the method evaluates molecular samples as monomers, regardless of their multimeric structure. This is the probable source of the discrepancy between the concentrations in Table 3 evaluated using the DNA assay and UV-Vis measurements. Alternatively, if the extinction coefficient for streptavidin monomers of 34,000 $M^{-1}cm^{-1}$ was used [43, 46] instead to evaluate the concentration by UV-Vis measurements, the same absorbance value would produce a concentration four times greater than using the tetramer extinction coefficient. However, since the mass concentration in the sample was the same, the four-fold lower molecular weight of the monomer must also be taken into account, which reduced the concentration by a factor of four, in agreement with our results. From these observations a significant benefit of the DNA assay emerges. It does not require prior knowledge of the sample to determine molecular stoichiometry.

It has also been reported that streptavidin is highly stable and even in the presence of SDS, temperatures above 60 °C are required to separate the tetramer into dimers and monomers [43]. Since SDS was not used here, even higher temperatures were required to dissociate the tetramer, consistent with other DSC experiments of streptavidin [47]. Once temperatures approach the T_m, the tetramer dissociates to four monomers, increasing the actual monomer concentration by a factor of four. The four identical monomers each melt with a corresponding equivalent transition

heat capacity. In this case the protein concentration evaluated from calorimetric measurements is approximately four times the tetramer concentration determined from absorbance measurements, and almost exactly what would be expected using the monomer extinction coefficient.

Another example was purified HSV-GpD. The evaluated concentration in column 4 of Table 3 was 1.06 mg/mL, a factor of 2.65 greater than the independently determined value of 0.40 mg/mL. Again, this finding is consistent with the known structure of HSV-GpD where there have been several reports that pure solutions of the protein contain both dimer (stoichiometry = 2) and trimer (stoichiometry = 3) forms [48-50]. In analogy to streptavidin, the ratio of 2.65 for HSV-GpD indicates that the protein solution was likely comprised of a mixture of dimer and trimer species, with slightly more trimer. This example provides additional support for the proposition that the DNA assay can also inform on the molecularity of multimeric proteins.

Likewise duplex DNA is a dimer comprised of two strands (stoichiometry = 2), whose base sequences are complementary. Concentrations determined by UV-Vis at 20°C assume a duplex (dimer) comprised of two monomer single strands. In analogy to streptavidin the duplex is assumed to be a "monomer" at 20°C, that then melts to two monomer single strands at higher temperatures. The duplex "monomer" concentration is twice the concentration of each single stand monomer. In this scenario the concentrations determined by UV-Vis at 20°C should be multiplied by the stoichiometry for duplex DNA.

As shown in Table 3 concentrations determined from the melting behavior of these molecules over the temperature range from 30 to 100°C, were 3.11 compared to 4 x 0.79= 3.16 mg/mL for streptavidin. For HSV-GpD containing a mixture of dimers and trimers compare 1.06 mg/mL to the range of 0.8-1.2 mg/mL. For duplex DNA compare 1.83 mg/mL to 2 x 0.85= 1.70 mg/mL. Results for these examples are all consistent with the notion that the higher $C_{p,PRO}$ and $C_{p,DNA}$ values, and higher concentrations determined from them, are indicative of higher stoichiometries of their tertiary structures, compared to the monomer. In effect, the DNA assay provides the monomeric concentration, regardless of oligomeric state. Thus, as demonstrated for the examples given above, the calorimetric method also provides insight into molecular tertiary structure where standard UV-Vis measurements cannot.

Proteins in Complex Solutions

These experiments involved evaluating the concentration of a single protein target in complex solutions, as opposed to a single protein in buffer alone.

Plasma: Human plasma served as an ideal test case. Over 9000 individual proteins comprise the plasma proteome, but a plasma thermogram is primarily due to seven major proteins [51, 52]. HSA comprises the major peak on a plasma thermogram. The height of this peak was used to estimate the fractional concentration of HSA in the complex mixture with other plasma protein components. As shown in Table 3 (row 5) there was good agreement between the reported fractional concentration of HSA in whole plasma 55-60% [22, 51-55], and the concentration of

11

HSA extracted from the peak height on the whole plasma thermogram, estimated to be 0.98 mg/mL, or 54.4% of the total concentration of the plasma, 1.8 mg/mL. In addition to HSA, six other proteins can also contribute to the peak height in a minor way. It should be noted that because their individual thermograms significantly overlap, the HSA peak on the plasma thermogram centered at ~62°C also contains slight contributions from the other six proteins most notedly alpha acid glycoprotein. However, since alpha acid glycoprotein (and the other proteins) comprise a small percentage (<1%) of the total concentration of human plasma, effects on the plasma thermogram compared to HSA are largely negligible [52].

E. coli **lysate:** To mimic conditions comparable to those possibly encountered for the unpurified lysates of expression systems, lyophilized *E. coli* cellular lysate was purchased from Bio-Rad, and a known concentration was re-suspended in standard NaCl buffer. The thermogram for the *E. coli* lysate displayed a broad curve from ~40-90°C with no significant distinguishing features [18]. The absolute height of this broad peak was used to estimate the concentration of lysate in the sample. Since the sample contains many unidentified proteins, a PSV value of 0.725 mL/g was used as the average of typical values reported for globular proteins (0.70-0.75 mL/mg) [32-36]. As Table 3 (row 6) shows, despite the complex milieu of the lysate, the peak height on the broad thermogram of the lysate alone, yielded a predicted concentration in agreement (within 10%) of the actual reported concentration of protein in the lysate solution. Although intriguing, this good agreement may be fortuitous and has not been independently substantiated.

HEK293 supernatant: To explore the more realistic environment of an actual expression system, unpurified ACE2 was purchased and used for comparison with a purified ACE2 sample. Unpurified ACE2 contained ACE2 expressed from HEK293 cell lines in cellular supernatant. The thermograms measured for the supernatant alone and for supernatant containing an unknown amount of expressed ACE2 are displayed in Fig 3a. Both curves were broad over the temperature range
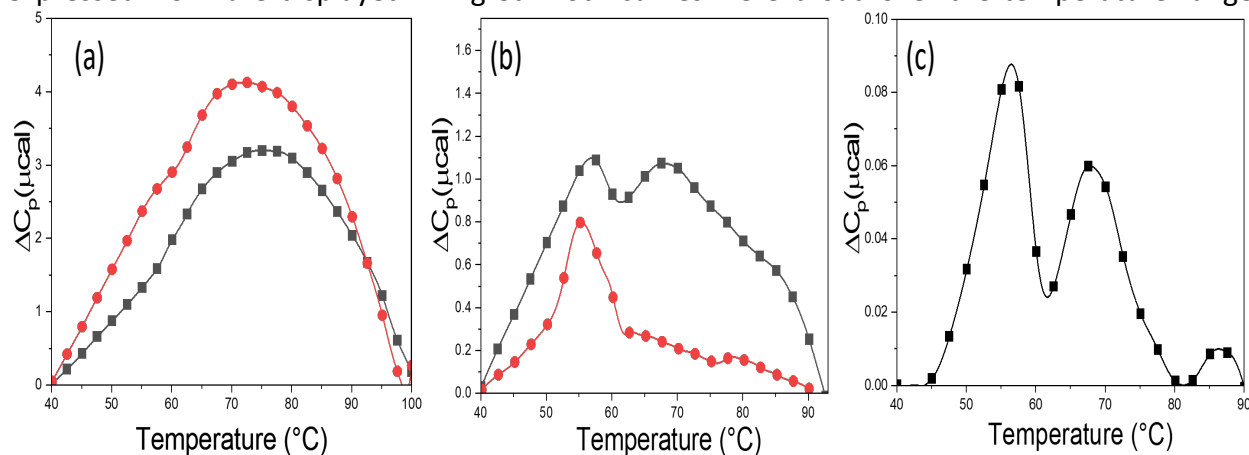

Figure 3: (a) Thermograms for the HEK293 cell supernatant background (black line) and a solution of the supernatant containing an unknown amount of expressed ACE2 (red line). (b) The difference thermogram corresponding to ACE2 protein alone obtained by subtracting the curves in (a) (black line); thermogram of a 0.20 mg/mL sample of purified ACE2 (red line). (c) Baseline corrected ACE2 curve from (b) and consistent with the reported thermogram for bovine somatic ACE2 [42].

displayed (40-100 ˚C), and the thermogram for ACE2 solution had a greater intensity for about 80% of the transition region. Subtracting the curves in Fig 3a provided the curve for ACE2 alone shown in Fig 3b. This curve clearly displayed two peaks around 55°C and 70°C. Also, shown in Fig 3b is the thermogram collected for purified ACE2 (200 μg/mL), that exhibits a single peak at ~55°C [18, 42]. Subtracting background baseline from the extracted curve for AC2 shown in Fig 3b resulted in the curve for unpurified AC2 shown in Fig 3c, where the two peaks are clearly displayed. There is clear agreement with published results [42]. The thermogram in Fig 3c is very similar to the published thermogram for somatic bovine ACE2 [42].

Baseline subtracted thermograms were used to estimate the concentration contributions from each domain of the ACE2 sample in Table 3 (row 7). For the C-domain, the concentration corresponding to the peak at ~55°C, was 0.04 mg/mL. Whereas for the N-domain, corresponding to the peak at ~70°C, the concentration was 0.03 mg/mL. Our method was able to provide an estimate for the amount of ACE2 in the unpurified cellular supernatant of the HEK293 expression system. As it was unknown to them, the vendor was unable to provide the actual concentration of ACE2 in the unpurified sample. Accuracies of these concentration estimates await verification.

DISCUSSION

Equivalance of Transition Heat Capacities. Affirming evidence was obtained for the equivalence of the transition heat capacites at T_m for proteins, $C_{p,PRO}$, and DNA, $C_{p,DNA}$ at the same mass concentration. Two DNA hairpins with a nearly 20°C difference in T_m values were examined in NaCl buffer and had the same $C_{p,DNA}$. Thermograms for these hairpins and HSA in CsCl buffer also provided identical results. It should be noted that the numerical values for the transition heat capacities, at the T_m, shown in Table 3 differ from the reported heat capacity at 25°C for HSA, lysozyme and other globular proteins [5, 19, 29]. Based on the equivalence of $C_{p,PRO}$ and $C_{p,DNA}$, several novel applications have emerged.

Concentration determination. A scheme was devised to determine unknown concentrations of individual proteins in pure and complex mixtures. Minimally, the scheme requires a single thermogram measured for both the protein of unknown concentration and DNA hairpin at a known concentration, along with the PSV values for DNA and proteins. Greater accuracy can be obtained if the entire DNA standard curve is employed, particularly at lower concentrations. The concentrations reported in Table 3 evaluated using the DNA standard curve agreed with concentrations determined by UV-Vis and dry weight measurements. It seems unlikely that consistent accuracy of evaluated concentrations, for such a wide variety of proteins, would have been possible if the transition heat capacities were not essentially equivalent. This was strictly true for all the monomer globular proteins examined (thus far). As discussed subsequently, additional evidence was provided suggesting the multimeric structures of streptavidin (a tetramer), HSV-GpD (dimer/trimer mix) and duplex DNA (a dimer) are also reflected in their transition heat capacities.

Multimeric Proteins. Examples of monomeric proteins that were analyzed included HSA, Lysozyme, and SARS-CoV-2 RBD. The tertiary structures of these proteins form from a single

polypeptide chain. Even though the tertiary structure of HSA contains three domains, the thermogram displays a single transition. Despite their ordered secondary and tertiary structures, the other proteins also displayed single peaks on their thermograms. In contrast, ACE2 is a heterodimer comprised of two different non-covalently associated monomeric subunits, the C- and N-domains [42, 56, 57]. These domains apparently differ enough in stability that the thermogram for ACE2 displayed two distinct peaks, presumably with each monomeric subunit displaying a distinguishable transition. For ACE2 the overall protein concentration in solution was expected to be 0.07 mg/mL. From the transition peak heights on the thermogram for ACE2 the evaluated concentrations were 0.04 mg/mL for the C-domain and 0.03 mg/mL of the N-domain for a total of 0.07 mg/mL for the total protein.

The values of the transition heat capacities found for streptavidin, HSV-GpD and linear duplex DNA were different from the other proteins and DNAs. Streptavidin is comprised of four identical protein monomers [44]. Unlike the other proteins that were examined, for streptavidin the evaluated concentration differed significantly from that determined from UV-Vis measurements. This concentration difference in Table 3 (row 8) was preserved over a range of concentrations and multiple experiments. Therefore, experimental error cannot account for the apparent discrepancy. The ratio of calorimetrically evaluated concentration to the UV-Vis determined concentration for the tetrameric protein was nearly fourfold higher (3.94) than observed for the monomeric proteins. Since all results were normalized by concentration, the observed fourfold difference in peak height was initially surprising but perhaps not unexpected considering the stoichiometric composition of streptavidin. In the case of streptavidin, the calorimetric method appears to be sensitive to not only contributions from separation of the four monomers in the overall secondary structure, but also melting of the four individual monomers comprising the tertiary structure. Likewise, for HSV-GpD the ratio of the evaluated concentration to the UV-Vis concentration was a factor of 2.65 suggesting a mixture of dimers and trimers at approximately, 35% and 65%, respectively.

A novel aspect of the calorimetric approach is the insights it can provide into tertiary structural composition of unknown proteins. This capability is not available with conventional spectroscopic methods such as UV-Vis and fluorescence. For example, as demonstrated here, calorimetric analysis was able to differentiate whether protein domains form from a contiguous primary structure i.e. monomeric (HSA, lysozyme); or are multimeric i.e. dimeric (unpurified ACE2), dimer and trimer mix (HSV-GpD) and tetrameric (streptavidin) formed by non-covalent association of identical monomer subunits. The calorimetric method evaluates the monomeric concentration, regardless of the multimeric state.

Duplex DNA: Thermograms collected for the 16 base-pair double stranded DNA as function of concentration are shown in Fig 4. Note, the curves incrementally shift to higher temperature as the concentration increases. Behavior entirely expected for the concentration dependence of short linear duplex DNA melting [58, 59]. In precisely the same manner as for the hairpin DNAs and protein samples the DNA concentration was determined from the peak heights on DSC

thermograms. The evaluated concentration was 2.15 greater than determined by UV-Vis. While one thermogram at a single concentration was used to evaluate the DNA concentration in Table 3, due to the linear relationship between the measured peak height and protein concentration, regardless of the concentration chosen, the resulting evaluated concentration was a factor of 2.15 greater than determined by UV-Vis. These results for short duplex DNAs are analogous to those found for streptavidin and HSV-GpD, and consistent with the bi-molecular composition of a DNA duplex dimer that forms via association of two independent single strand monomers. Although these are reasonable and compelling scenarios for streptavidin and HSV-GpD, further calorimetric examination of additional multimeric proteins and multi-stranded DNAs will be required before these interpretations can be generalized.
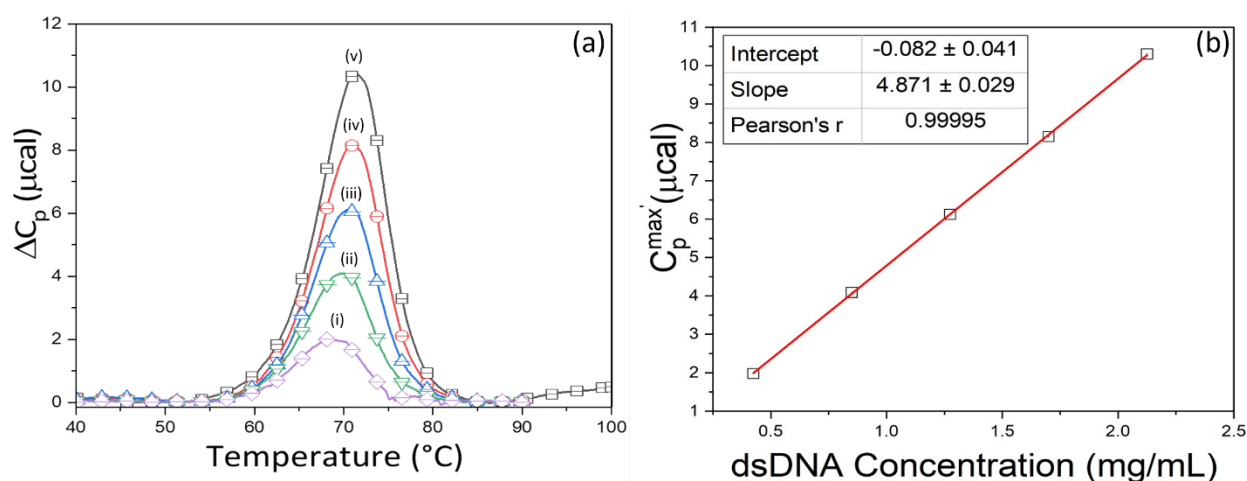

Figure 4: (a) Average thermograms for dsDNA concentrations in mg/mL: (i) 0.43, (ii) 0.85, (iii) 1.28, (iv) 1.70, (v) 2.13. (b) Transition heat capacity curve for dsDNA.

A tool for Drug Discovery and Screening. Insight provided by the observed equivalence of the transition heat capacities for proteins and DNA provides for novel modern applications. For example, *in vitro* expression systems (prokaryotic and eukaryotic) are the source of many specially engineered low abundance molecules and biopharmaceuticals (biologics) [60]. Expression systems provide a means for generating important biological molecules (proteins, peptides, antibodies), on various scales, relatively fast for reasonable cost. Unfortunately, isolation and purification of expressed molecules can be expensive and time-consuming. For a moderately abundant and soluble protein, conventional purification strategies can require 27 or more individual steps and at least four days to complete [61]. In some cases, additional time-consuming purification, fractionation, and screening procedures may be required to reach desired levels of purity. This time requirement currently constitutes an obstacle to biologic drug development and emphasizes the need for a quick and quantitatively reliable means to reliably estimate unknown concentrations of protein products of *in vitro* over-expression systems. In principle, the demonstrated capability of calorimetric analysis could provide a means for quickly

screening libraries of potential compounds for yield and activity, without requiring prior purification. In so doing, relieving a critical bottleneck in biopharmaceutical development.

Previously we demonstrated the utility of DSC to detect and analyze ligand binding to unpurified protein products of expression systems [18]. At that time, it was not possible to quantitatively evaluate ligand binding constants for proteins whose concentrations were unknown. As reported, this was the case for an unpurified protein (ACE 2) product of an expression system [18]. Now that we have demonstrated the ability to quantitatively evaluate the concentration of specific proteins in complex media mixtures, it should be possible to use DSC to quantitatively evaluate ligand binding constants to protein products of expression systems.

DSC versus UV-Vis. The reported calorimetric DNA assay provides unique information that other methods such as absorbance measurements (UV-Vis) cannot. As has been demonstrated, the calorimetric scheme enables quantitative evaluation of unknown protein concentrations in pure solutions and complex mixtures, without the need for probes or binding assays. Insights into molecular stoichiometry of proteins and DNA can also be gleaned from calorimetric measurements. An added feature of calorimetric analysis is the much higher solution concentrations that can be analyzed compared to UV-Vis absorbance measurements.

UV-Vis measurements can be used to quantitatively evaluate concentrations of biological molecules. Practically, UV-Vis measurements require samples having sufficient optical activity in the UV, with measurable spectroscopic properties or the ability to bind a colorimetric probe. UV-Vis most commonly requires solutions, in a one cm pathlength cuvette, having UV absorbance ≤ 1 OD/ml. For more concentrated solutions, accurate absorbance measurements require dilutions to achieve readings ≤ 1 OD/ml. Additional sample handling involved in preparing diluted samples can also be an unnecessary source of error.

One practical difference and consequence of calorimetric analysis is that it requires melting of the protein sample. During the melting process, typically over the range from 30 to 90°C, the sample may get damaged, or irreversibly aggregate. Consequently, the sample may not be retrievable in active form after the experiment and consequently sacrificed. This is not the case for UV-Vis measurements at 20°C where the sample can usually be safely retrieved.

Conclusions. Precise origins of the observed heat capacity equivalence for proteins and DNA remain a mystery, but none-the-less an empirical fact for the examples examined thus far. Several potential sources are merely mentioned here. The suspected role of hydration has been discussed and likely a prevalent component [14, 19, 32, 62-68]. In addition to hydration, electrostatic and vibrational effects can also contribute to the transition heat capacity [69, 70]. Our results in CsCl indicated no differences for the DNAs and HSA from what was found in standard NaCl buffer. Cesium has a much larger (50%) Van der Waals radius and 5.75 times greater mass (132.91 g/mol) than sodium (22.99 g/mol). If solvent interactions contributed significantly to $C_{p,DNA}$ and $C_{p,PRO}$ then different values in NaCl versus CsCl might have been expected. None were observed suggesting electrostatics and associated polyelectrolyte effects

on the observed equivalence must be relatively small. This is in accordance with earlier modeling studies suggesting electrostatics only contribute in a minor way to the heat capacity [69].

Several examples of our novel approach to calorimetric measurements and analysis have been previously reported [17-19]. As described here the DNA assay for evaluating concentrations of protein solutions from DSC thermograms is the most recent of these applications. Others have involved using calorimetric analysis to examine ligand binding reactions in both pure solutions and complex media mixtures [18]. The essence of our approach lies in the power of relative measurements which have been shown to provide a sensitive and accurate means to examine protein structure and ligand binding. For example, *ratiometric* analyses of DSC thermograms, versus absolute measurements, are much more accurate and informative, without the added requirement of tenuous model building, associated underlying assumptions and analysis [17-19].

Clearly the calorimetric approach provides unique capabilities with potentially powerful applications. With the novel information that can be obtained using calorimetry important applications in drug development and molecular diagnostics seem obvious. Unfortunately, these applications of DSC are currently not realistic simply because of the serial nature of sample processing with standard bio calorimeters. i.e. samples can only be processed one at a time per instrument. An array of instruments is logistically impractical. So, realization of the untapped potential of calorimetry in important prescient applications will first require a new generation of DSC instruments. To meet the demands for high-throughput applications these next generation calorimeters must be built with the necessary signal sensitivity and have multiplex capabilities of processing multiple samples simultaneously in parallel.

Data Availability

The data underlying this article are available in the article and in its online supplementary material.

Acknowledgements

References

[1]   J. M. Walker, "The bicinchoninic acid (BCA) assay for protein quantitation," *The protein protocols handbook,* pp. 11-15, 2009.

[2]   L. J. Jones, R. P. Haugland, and V. L. Singer, "Development and characterization of the NanoOrange® protein quantitation assay: a fluorescence-based assay of proteins in solution," *Biotechniques,* vol. 34, no. 4, pp. 850-861, 2003.

[3]   J. E. Noble and M. J. Bailey, "Quantitation of protein," *Methods in enzymology,* vol. 463, pp. 73-95, 2009.

[4]   N. J. Kruger, "The Bradford method for protein quantitation," *The protein protocols handbook,* pp. 17-24, 2009.

[5]   E. Freire, "Differential scanning calorimetry," (in eng), *Methods Mol Biol,* vol. 40, pp. 191-218, 1995, doi: 10.1385/0-89603-301-5:191.

[6]   J. M. Sanchez-Ruiz, "Differential scanning calorimetry of proteins," in *Proteins: Structure, Function, and Engineering*: Springer, 1995, pp. 133-176.

[7]   P. V. Riccelli, K. E. Mandell, and A. S. Benight, "Melting studies of dangling-ended DNA hairpins: effects of end length, loop sequence and biotinylation of loop bases," *Nucleic Acids Research,* vol. 30, no. 18, pp. 4088-4093, 2002, doi: 10.1093/nar/gkf514.

[8]   G. Bruylants, J. Wouters, and C. Michaux, "Differential scanning calorimetry in life science: thermodynamics, stability, molecular recognition and application in drug design," (in eng), *Curr Med Chem,* vol. 12, no. 17, pp. 2011-20, 2005.

[9]   G. Holdgate, "Isothermal titration calorimetry and differential scanning calorimetry," (in eng), *Methods Mol Biol,* vol. 572, pp. 101-33, 2009, doi: 10.1007/978-1-60761-244-5_7.

[10]   R. Splinter, A. W. van Herwaarden, E. Iervolino, G. V. Poel, D. Istrate, and P. M. Sarro, "Analyzing Protein Denaturation using Fast Differential Scanning Calorimetry," *Procedia Engineering,* vol. 47, no. Supplement C, pp. 140-143, 2012/01/01/ 2012, doi: https://doi.org/10.1016/j.proeng.2012.09.104.

[11]   C. M. Johnson, "Differential scanning calorimetry as a tool for protein folding and stability," (in eng), *Arch Biochem Biophys,* vol. 531, no. 1-2, pp. 100-9, Mar 2013, doi: 10.1016/j.abb.2012.09.008.

[12]   C. M. Johnson, "Differential scanning calorimetry as a tool for protein folding and stability," *Archives of biochemistry and biophysics,* vol. 531, no. 1-2, pp. 100-109, 2013.

[13]   N. C. Garbett and G. N. Brock, "Differential scanning calorimetry as a complementary diagnostic tool for the evaluation of biological samples," (in eng), *Biochim Biophys Acta,* vol. 1860, no. 5, pp. 981-989, May 2016, doi: 10.1016/j.bbagen.2015.10.004.

[14]   A. Cooper, C. M. Johnson, J. H. Lakey, and M. Nöllmann, "Heat does not come in different colours: entropy–enthalpy compensation, free energy windows, quantum confinement, pressure perturbation calorimetry, solvation and the multiple causes of heat capacity effects in biomolecular interactions," *Biophysical chemistry,* vol. 93, no. 2-3, pp. 215-230, 2001.

[15]   G. Mrevlishvili, A. Carvalho, M. R. da Silva, T. Mdzinarashvili, G. Razmadze, and T. Tarielashvili, "The role of bound water on the energetics of DNA duplex melting," *Journal of thermal analysis and calorimetry,* vol. 66, no. 1, pp. 133-144, 2001.

[16]   M. W. Eskew, M. M. Koslen, and A. S. Benight, "Ligand binding to natural and modified human serum albumin," *Analytical Biochemistry,* vol. 612, no. 1, p. 113843, 2021/01/01/ 2021, doi: https://doi.org/10.1016/j.ab.2020.113843.

[17]   M. W. Eskew and A. S. Benight, "Ligand binding constants for human serum albumin evaluated by ratiometric analysis of DSC thermograms," *Analytical Biochemistry,* vol. 628, p. 114293, 2021.

[18]   M. W. Eskew and A. S. Benight, "Thermal analysis of protein stability and ligand binding in complex media," *Thermochimica Acta,* vol. 715, p. 179274, 2022.

[19]   M. W. Eskew and A. S. Benight, "Equivalence of the transition heat capacities of proteins and DNA," *Biochemical and Biophysical Research Communications,* vol. 597, pp. 98-101, 2022.

[20]   A. S. Benight *et al.*, "Conformation and dynamics of a left-handed Z-DNA hairpin: Studies of d (CGCGCGTTTTCGCGCG) in solution," *Biochemistry,* vol. 28, no. 8, pp. 3323-3332, 1989.

[21]   P. M. Vallone, T. M. Paner, J. Hilario, M. J. Lane, B. D. Faldasz, and A. S. Benight, "Melting studies of short DNA hairpins: influence of loop sequence and adjoining base pair identity on hairpin thermodynamic stability," *Biopolymers: Original Research on Biomolecules,* vol. 50, no. 4, pp. 425-442, 1999.

[22]   T. Peters Jr, *All about albumin: biochemistry, genetics, and medical applications*. Academic press, 1995.

[23]   A. Beaussart *et al.*, "Supported lysozyme for improved antimicrobial surface protection," *Journal of Colloid and Interface Science,* vol. 582, pp. 764-772, 2021.

[24]   H. Hoang *et al.*, "Effects of Selective Biotinylation on the Thermodynamic Stability of Human Serum Albumin," *Journal of Biophysical Chemistry,* vol. 7, no. 01, p. 9, 2016.

[25]   M. M. Koslen *et al.*, "Capture Reagent and Strategy for Retrieving Albumin-Bound Ligands from Plasma," *Advances in Biological Chemistry,* vol. 9, no. 3, 2019.

[26]   K. Pearson, "Determination of the coefficient of correlation," *Science,* vol. 30, no. 757, pp. 23-25, 1909.

[27]   A. G. Asuero, A. Sayago, and A. González, "The correlation coefficient: An overview," *Critical reviews in analytical chemistry,* vol. 36, no. 1, pp. 41-59, 2006.

[28]   J. L. Cole, J. W. Lary, T. P. Moody, and T. M. Laue, "Analytical ultracentrifugation: sedimentation velocity and sedimentation equilibrium," *Methods in cell biology,* vol. 84, pp. 143-179, 2008.

[29]   V. Kholodenko and E. Freire, "A simple method to measure the absolute heat capacity of proteins," *Analytical biochemistry,* vol. 270, no. 2, pp. 336-338, 1999.

[30]   H. T. Le, R. Buscaglia, W. L. Dean, J. B. Chaires, and J. O. Trent, "Calculation of hydrodynamic properties for G-quadruplex nucleic acid structures from in silico bead models," *Quadruplex Nucleic Acids,* pp. 179-210, 2012.

[31]   G. Cohen and H. Eisenberg, "Deoxyribonueleate solutions: Sedimentation in a density gradient, partial specific volumes, density and refractive index increments, and preferential interactions," *Biopolymers: Original Research on Biomolecules,* vol. 6, no. 8, pp. 1077-1100, 1968.

[32]   L. R. Murphy, N. Matubayasi, V. A. Payne, and R. M. Levy, "Protein hydration and unfolding–insights from experimental partial specific volumes and unfolded protein models," *Folding and Design,* vol. 3, no. 2, pp. 105-118, 1998.

[33]   M. Young, P. Carroad, and R. Bell, "Estimation of diffusion coefficients of proteins," *Biotechnology and bioengineering,* vol. 22, no. 5, pp. 947-955, 1980.

[34]   H. P. Erickson, "Size and shape of protein molecules at the nanometer level determined by sedimentation, gel filtration, and electron microscopy," *Biological procedures online,* vol. 11, no. 1, pp. 32-51, 2009.

[35]   D. M. Hatters, L. Wilson, B. W. Atcliffe, T. D. Mulhern, N. Guzzo-Pernell, and G. J. Howlett, "Sedimentation analysis of novel DNA structures formed by homo-oligonucleotides," *Biophysical journal,* vol. 81, no. 1, pp. 371-381, 2001.

[36]   L. M. Hellman, D. W. Rodgers, and M. G. Fried, "Phenomenological partial-specific volumes for G-quadruplex DNAs," *European Biophysics Journal,* vol. 39, no. 3, pp. 389-396, 2010.

[37]   F. Li, M. Berardi, W. Li, M. Farzan, P. R. Dormitzer, and S. C. Harrison, "Conformational states of the severe acute respiratory syndrome coronavirus spike protein ectodomain," *Journal of virology,* vol. 80, no. 14, pp. 6794-6800, 2006.

[38] H. Zhao *et al.*, "Plasticity in structure and assembly of SARS-CoV-2 nucleocapsid protein," *PNAS nexus,* vol. 1, no. 2, p. pgac049, 2022.

[39] F. Li, W. Li, M. Farzan, and S. C. Harrison, "Structure of SARS coronavirus spike receptor-binding domain complexed with receptor," *Science,* vol. 309, no. 5742, pp. 1864-1868, 2005.

[40] A. Pähler, W. A. Hendrickson, M. Kolks, C. Argarana, and C. R. Cantor, "Characterization and crystallization of core streptavidin," *Journal of Biological Chemistry,* vol. 262, no. 29, pp. 13933-13937, 1987.

[41] C.-C. Lee, L.-L. Lin, W.-E. Chan, T.-P. Ko, J.-S. Lai, and A.-J. Wang, "Structural basis for the antibody neutralization of herpes simplex virus," *Acta Crystallographica Section D: Biological Crystallography,* vol. 69, no. 10, pp. 1935-1945, 2013.

[42] S. Voronov, N. Zueva, V. Orlov, A. Arutyunyan, and O. Kost, "Temperature-induced selective death of the C-domain within angiotensin-converting enzyme molecule," *FEBS letters,* vol. 522, no. 1-3, pp. 77-82, 2002.

[43] E. A. Bayer, H. Ben-Hur, and M. Wilchek, "[8] Isolation and properties of streptavidin," in *Methods in enzymology*, vol. 184: Elsevier, 1990, pp. 80-89.

[44] R. E. Stenkamp, I. L. Trong, L. Klumb, P. S. Stayton, and S. Freitag, "Structural studies of the streptavidin binding loop," *Protein Science,* vol. 6, no. 6, pp. 1157-1166, 1997.

[45] N. Malmstadt, D. E. Hyre, Z. Ding, A. S. Hoffman, and P. S. Stayton, "Affinity Thermoprecipitation and Recovery of Biotinylated Biomolecules via a Mutant Streptavidin– Smart Polymer Conjugate," *Bioconjugate chemistry,* vol. 14, no. 3, pp. 575-580, 2003.

[46] N. M. Green, "Avidin," in *Advances in protein chemistry*, vol. 29: Elsevier, 1975, pp. 85-133.

[47] M. González *et al.*, "Interaction of biotin with streptavidin thermostability and conformational changes upon binding," *Journal of Biological Chemistry,* vol. 272, no. 17, pp. 11288-11294, 1997.

[48] S. H. Willis *et al.*, "Examination of the kinetics of herpes simplex virus glycoprotein D binding to the herpesvirus entry mediator, using surface plasmon resonance," *Journal of virology,* vol. 72, no. 7, pp. 5937-5947, 1998.

[49] R. Eisenberg, M. Ponce de Leon, L. Pereira, D. Long, and G. Cohen, "Purification of glycoprotein gD of herpes simplex virus types 1 and 2 by use of monoclonal antibody," *Journal of Virology,* vol. 41, no. 3, pp. 1099-1104, 1982.

[50] C. G. Handler, R. J. Eisenberg, and G. H. Cohen, "Oligomeric structure of glycoproteins in herpes simplex virus type 1," *Journal of Virology,* vol. 70, no. 9, pp. 6067-6070, 1996.

[51] N. C. Garbett, J. J. Miller, A. B. Jenson, and J. B. Chaires, "Calorimetric analysis of the plasma proteome," (in eng), *Semin Nephrol,* vol. 27, no. 6, pp. 621-6, Nov 2007, doi: 10.1016/j.semnephrol.2007.09.004.

[52] N. C. Garbett, J. J. Miller, A. B. Jenson, and J. B. Chaires, "Calorimetry outside the box: a new window into the plasma proteome," (in eng), *Biophys J,* vol. 94, no. 4, pp. 1377-83, Feb 2008, doi: 10.1529/biophysj.107.119453.

[53] N. L. Anderson *et al.*, "The human plasma proteome: a nonredundant list developed by combination of four separate sources," (in eng), *Mol Cell Proteomics,* vol. 3, no. 4, pp. 311-26, Apr 2004, doi: 10.1074/mcp.M300127-MCP200.

[54] G. S. Omenn *et al.*, "Overview of the HUPO Plasma Proteome Project: results from the pilot phase with 35 collaborating laboratories and multiple analytical groups, generating a core dataset of 3020 proteins and a publicly-available database," (in eng), *Proteomics,* vol. 5, no. 13, pp. 3226-45, Aug 2005, doi: 10.1002/pmic.200500358.

[55] N. L. Anderson and N. G. Anderson, "The human plasma proteome: history, character, and diagnostic prospects," (in eng), *Mol Cell Proteomics,* vol. 1, no. 11, pp. 845-67, Nov 2002.

[56]   P. Towler *et al.*, "ACE2 X-ray structures reveal a large hinge-bending motion important for inhibitor binding and catalysis," *Journal of Biological Chemistry,* vol. 279, no. 17, pp. 17996-18007, 2004.

[57]   P. Prabakaran, X. Xiao, and D. S. Dimitrov, "A model of the ACE2 structure and function as a SARS-CoV receptor," *Biochemical and biophysical research communications,* vol. 314, no. 1, pp. 235-241, 2004.

[58]   R. Owczarzy, P. M. Vallone, F. J. Gallo, T. M. Paner, M. J. Lane, and A. S. Benight, "Predicting sequence-dependent melting stability of short duplex DNA oligomers," *Biopolymers: Original Research on Biomolecules,* vol. 44, no. 3, pp. 217-239, 1997.

[59]   A. S. Benight, P. Pančoška, R. Owczarzy, P. M. Vallone, J. Nešetřil, and P. V. Riccelli, "Calculating sequence-dependent melting stability of duplex DNA oligomers and multiplex sequence analysis by graphs," *Methods in enzymology,* vol. 340, pp. 165-192, 2001.

[60]   D. C. Andersen and L. Krummen, "Recombinant protein expression for therapeutic applications," *Current opinion in biotechnology,* vol. 13, no. 2, pp. 117-123, 2002.

[61]   P. T. Wingfield, "Preparation of soluble proteins from Escherichia coli," *Current protocols in protein science,* vol. 78, no. 1, pp. 6.2. 1-6.2. 22, 2014.

[62]   V. A. Sirotkin, I. A. Komissarov, and A. V. Khadiullina, "Hydration of proteins: excess partial volumes of water and proteins," *The Journal of Physical Chemistry B,* vol. 116, no. 13, pp. 4098-4105, 2012.

[63]   G. W. Robinson and C. Cho, "Role of hydration water in protein unfolding," *Biophysical Journal,* vol. 77, no. 6, pp. 3311-3318, 1999.

[64]   O. Rahaman, M. Kalimeri, S. Melchionna, J. Hénin, and F. Sterpone, "Role of internal water on protein thermal stability: The case of homologous G domains," *The Journal of Physical Chemistry B,* vol. 119, no. 29, pp. 8939-8949, 2015.

[65]   P. L. Privalov and C. Crane-Robinson, "Role of water in the formation of macromolecular structures," *European biophysics journal,* vol. 46, no. 3, pp. 203-224, 2017.

[66]   T. Phan-Xuan *et al.*, "The role of water in the reversibility of thermal denaturation of lysozyme in solid and liquid states," *Biochemistry and Biophysics Reports,* vol. 28, p. 101184, 2021.

[67]   M. Adrover *et al.*, "The role of hydration in protein stability: comparison of the cold and heat unfolded states of Yfh1," *Journal of molecular biology,* vol. 417, no. 5, pp. 413-424, 2012.

[68]   F. Mallamace *et al.*, "Energy landscape in protein folding and unfolding," *Proceedings of the National Academy of Sciences,* vol. 113, no. 12, pp. 3159-3163, 2016.

[69]   B. Madan and K. A. Sharp, "Hydration heat capacity of nucleic acid constituents determined from the random network model," *Biophysical journal,* vol. 81, no. 4, pp. 1881-1887, 2001.

[70]   I. Rouzina and V. A. Bloomfield, "Heat capacity effects on the melting of DNA. 1. General aspects," *Biophysical journal,* vol. 77, no. 6, pp. 3242-3251, 1999.