Supplementary Information

Rapid characterization of folding and binding interactions with thermolabile ligands by DSC

Robert W. Harkness, V[‡], Sladjana Slavkovic[†], Philip E. Johnson[†], Anthony K. Mittermaier[‡]

McGill University, Department of Chemistry[‡]

801 Sherbrooke St. W., Montreal QC Canada

H3A 0B8

T: 514-398-2093

F: 514-398-3797

York University, Department of Chemistry[†]

4700 Keele Street, Toronto ON Canada

M3J 1P3

Supplementary Methods

DSC global fitting

Heat capacity profiles were analyzed assuming unfolded (U), folded (F), folded cocaine-bound (B1), and folded benzoylecgonine-bound (B2) states in equilibrium

$$B2 \; \overset{K^{B2F}}{\rightleftharpoons} \; L2 + F + L1 \; \overset{K^{B1F}}{\rightleftharpoons} \; B1$$
$$K^{UF} \updownarrow$$
$$U \tag{1}$$
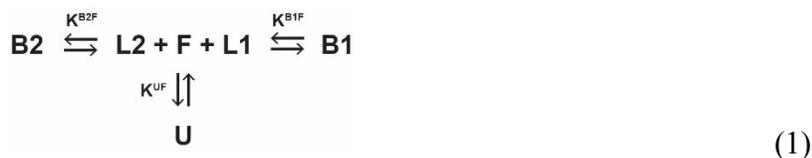
where L1 and L2 are cocaine and benzoylecgonine ligands respectively. The heat capacity profiles were fit using temperature dependent thermodynamic parameters for the ligand binding processes

$$\Delta H^{B1F}(T) = \Delta H^{B1F}(T_0) + \Delta C_p^{B1F}(T - T_0) \tag{2}$$

$$\Delta S^{B1F}(T) = \Delta S^{B1F}(T_0) + \Delta C_p^{B1F} \ln\left\{\frac{T}{T_0}\right\} \tag{3}$$

where $\Delta H^{B1F}(T)$, $\Delta S^{B1F}(T)$, and ΔC_p^{B1F} are the changes in enthalpies, entropies, and heat capacities for the folded to cocaine-bound (B1F) processes respectively and T_0 is the reference temperature. Folded to benzoylecgonine-bound (B2F) parameters were fit using the same equations as the B1F equilibrium but these have been omitted here for the sake of brevity. The change in heat capacity for unfolding, ΔC_p^{UF}, was set equal to zero as it was found to be negligibly small when included in the global fits. Equilibrium constants for the folding and binding processes were calculated according to

$$K^{UF}(T) = \frac{[U](T)}{[F](T)} = \exp\left(\frac{-\left(\Delta H^{UF} - T\Delta S^{UF}\right)}{RT}\right) \tag{4}$$

$$K^{B1F}(T) = \frac{[B1](T)}{[F](T)[L1](T)} = \exp\left(\frac{-\left(\Delta H^{B1F}(T) - T\Delta S^{B1F}(T)\right)}{RT}\right) \tag{5}$$

and combined in the partition function assuming folded, cocaine-bound, benzoylecgonine-bound, and unfolded states according to

$$Z(T) = 1 + K^{UF}(T) + K^{B1F}(T)[L1](T) + K^{B2F}(T)[L2](T) \tag{6}$$

where 1 is the relative population of the folded state. The concentration of folded state as a function of temperature (or similarly the unfolded state for MN19 as [F](T) = 0) was obtained by numerically solving the real, positive root of

$$a[F](T)^3 + b[F](T)^2 + c[F](T) - [C]_T = 0 \tag{7}$$

where $a = K^{UF}(T)\, K^{B1F}(T)\, K^{B2F}(T) + K^{B1F}(T)\, K^{B2F}(T)$, $b = K^{UF}(T)\, K^{B1F}(T) + K^{UF}(T)\, K^{B2F}(T) + K^{B1F}(T) + K^{B2F}(T) - [C]_T K^{B1F}(T)\, K^{B2F}(T) + [L2]_T K^{B1F}(T)\, K^{B2F}(T) + [L1]_T K^{B1F}(T)\, K^{B2F}(T)$, and $c = 1 + K^{UF} - [C]_T K^{B1F} - [C]_T K^{B2F} + [L1]_T K^{B1F}(T) + [L2]_T K^{B2F}(T)$, and $[C]_T$, $[L1]_T$, and $[L2]_T$ are the total concentrations of aptamer, cocaine, and benzoylecgonine respectively. Free ligand concentrations $[L1](T)$ and $[L2](T)$ were calculated using

$$[L1](T) = \frac{[L1]_T}{1 + K^{B1F}[F]}$$

(8)

And the populations of each state were computed as

$$P^U(T) = \frac{K^{UF}(T)}{Z(T)} \tag{9}$$

$$P^{B1}(T) = \frac{K^{B1F}(T)[L1](T)}{Z(T)} \tag{10}$$

from which the DSC thermogram (C_p) profiles were calculated

$$
\begin{aligned}
C_p^{calc}(T) &= C_p^F(T) + \Delta H^{UF} \times \frac{d}{dT} P^U(T) + \Delta H^{B1F}(T) \times \frac{d}{dT} P^{B1}(T) \\
&+ \Delta H^{B2F}(T) \times \frac{d}{dT} P^{B2}(T) + P^{B1}(T) \times \Delta C_p^{B1F} + P^{B2}(T) \times \Delta C_p^{B2F}
\end{aligned}
\tag{11}
$$

where $C_p^F(T)$ is the heat capacity of the reference folded state, calculated as a second order polynomial in temperature. Note that since ΔC_p^{UF} is zero, the heat capacity of the unfolded state is identical. The calculated thermograms were globally fit to the experimental DSC profiles by minimizing the residual sum-of-squared-differences (RSS)

$$RSS(\xi) = \sum_{i=1}^{M} \sum_{j=1}^{N} (C_p^{exp}(i,j) - C_p^{calc}(\xi, i, j))^2 \tag{12}$$

where M is the number of replicate DSC scans, N is the number of data points in each scan, and $\xi = [\Delta H^{UF}, \Delta H^{B1F}, \Delta H^{B2F}\ \Delta S^{UF}, \Delta S^{B1F}, \Delta S^{B2F}, \Delta C_p^{B1F}, \Delta C_p^{B2F}, [L1]_{T,i}, a_f, b_f, c_f]$. $(\Delta H^{UF}, \Delta H^{B1F}, \Delta H^{B2F}\ \Delta S^{UF}, \Delta S^{B1F}, \Delta S^{B2F}, \Delta C_p^{B1F}, \Delta C_p^{B2F})$ are the thermodynamic parameters for folding and binding, (a_f, b_f, c_f) define the C_p^F baseline, and $[L1]_{T,i}$ is the total cocaine concentration of the ith DSC profile. $[L1]_{T,1}$ was fixed at the initial ligand concentration and $[L1]_{T,2-M}$ were optimized in the fits as adjustable parameters. The total concentration of benzoylecgonine in each scan was calculated as $[L2]_{T,i} = [L1]_{T,0} - [L1]_{T,i}$, and the final scan of the cocaine-bound DSC manifold assumed cocaine conversion was 100% complete, i.e. $[L2]_{T,M} = [L1]_{T,0}$. The quinine-bound MN4 global fits were constrained by including the unbound dataset as $i=0$, with $L_{T,0}=0$.

Testing the high temperature ligand conversion assumption

In our global fits we used constant total ligand concentrations for each thermogram, based on the assumption that ligand thermal conversion predominantly occurs during the high temperature equilibration period. However, this assumption depends on both the temperature scan rate and the rate of ligand conversion. It is possible for the assumption to be violated with very slow temperature scanning or rapid ligand conversion. Slow scan rates lengthen the amount of time the ligand spends at each temperature and rapid ligand conversion causes the initial ligand to be depleted within the first scan. In our case, cocaine conversion is strongly temperature dependent and scanning at 1 °C min⁻¹ does not violate our assumption that most of the conversion happens at high temperatures. Simulations with continuously-varying cocaine concentrations at scan rates of 1 °C min⁻¹ are superimposable with those using fixed concentrations (Supplementary Figure 3a). We have simulated experiments where we varied the scan rate or the ligand conversion kinetics (by modifying the activation energy) in order to provide visual evidence of when the assumption breaks down (Supplementary Figure 3). The rate constants for ligand conversion were calculated every 0.1 °C from 20-80 °C using

$$k(T) = Ae^{-\frac{E_a}{RT}} \tag{13}$$

Where A is the pre-exponential factor (7.51×10^{10} s⁻¹ for cocaine), E_a is the activation energy for conversion (95.9 kJ mol⁻¹ for cocaine), and R is the ideal gas constant. Ligand concentrations at every 0.1 °C were computed with

$$[L]_t = [L]_{t-1}e^{-k(T) \times t} \tag{14}$$

Where $[L]_t$ is the new ligand concentration after converting for a time t (set by the scan rate, for example the average time per 0.1 °C at 1 °C min⁻¹ = 6 seconds) at temperature T, and $[L]_{t-1}$ is the ligand concentration at the previous temperature.

Clear distortions of the DSC peak shape occur when the ligand is depleted in the early portion of the thermogram, either when scanning extremely slowly or when rapid conversion occurs at lower temperatures. We note that the 0.005 °C min⁻¹ scan rate is unfeasible in practice as the DSC signal to noise becomes poor below scan rates of 0.1 °C min⁻¹. We found that a conversion ratio (CR, °C) defined as the scan rate (°C min⁻¹) divided by the rate constant for ligand conversion at the apparent T_m of the first forward scan (min⁻¹) gives a measure of when our ligand conversion assumption is violated. For conversion ratios below ~20 °C, substantial depletion of the ligand occurs before and during the thermogram leading to distortions of the transition shape. This indicates continuously-varying ligand concentrations must be applied in the fit. By increasing the scan rate, one may adjust the conversion ratio for a thermolabile ligand DSC experiment in order to obtain data that can be fit with constant ligand concentrations. The scan rate must however remain slow enough to avoid thermal hysteresis. This can be tested with an experiment on the free biomolecule.

Ligand conversion can also be modulated by protection of the ligand in the biomolecular binding pocket. The experiments were performed with ligand concentrations in excess of the biomolecule (10:1), i.e. the total amount of available ligand is largely unbound. Therefore, the overall ligand conversion is dominated by that of the free ligand molecules. Assuming the biomolecule can protect the ligand and the ligand is in excess, the apparent rate constant is given by $k^{app.}(T) = k(T) \times P_{free}(T)$, where $P_{free}(T)$ is the population of free ligand as a function of temperature. This assumes the equilibration between free and bound ligand is rapid relative to the conversion rate. Ligand concentrations at each temperature are accordingly calculated with Eq. 14. We have simulated the DSC profiles where protection of the ligand occurs at scan rates of 1 °C min⁻¹ and 0.005 °C min⁻¹, overlaying these with the case where ligand is not protected by the biomolecule and continuously varies at the corresponding scan rates (Supplementary Figure 4). We find that protection of the ligand does not appreciably modify the result, even at 0.005 °C min⁻¹ scan rate where it can play a greater role.

Calculation of the rate constant for conversion of cocaine to benzoylecgonine

The concentration of cocaine in any scan, C_N, is related to the concentration remaining in the subsequent scan, C_{N+1} (Supplementary Scheme 1), according to

$$C_{N+1} = f * C_N e^{-kt_{equil.}} \tag{15}$$

where t_{equil} is the length of the high-temperature (80 °C) equilibration period separating each heating scan from the following cooling scan, k is the rate constant for thermal conversion, and f (<1) accounts for the thermal conversion occurring during a cooling scan and subsequent heating scan. The factor f cancels out when a comparison is made between the ratio of two N scans and two $N+1$ scans obtained with different equilibration times (A and B), as follows:

$$\frac{C_{N+1}^B}{C_{N+1}^A} = \frac{f * C_N^B e^{-kt_{equil.}^B}}{f * C_N^A e^{-kt_{equil.}^A}} = \frac{C_N^B}{C_N^A} e^{-k\Delta t_{equil.}^{BA}} \tag{16}$$

And

$$k = -\ln\left(\frac{\left(\frac{C_{N+1}^B}{C_{N+1}^A}\right)}{\left(\frac{C_N^B}{C_N^A}\right)}\right)\Delta t_{equil.}^{BA}{}^{-1}. \tag{17}$$

Extending this to the general case for the initial forward scan (number *1*) and the *N*th later scan (number *1+N*) gives

$$\frac{C_{1+N}^B}{C_{1+N}^A} = \frac{f^N * C_1^B e^{-kNt_{equil.}^B}}{f^N * C_1^A e^{-kNt_{equil.}^A}} = \frac{C_1^B}{C_1^A} e^{-kN\Delta t_{equil.}^{BA}} \tag{18}$$

The rate constant for ligand thermal conversion can be obtained from fitting

$$\ln\left(\frac{\left(\dfrac{C^B_{1+N}}{C^A_{1+N}}\right)}{\left(\dfrac{C^B_1}{C^A_1}\right)}\right) = -kN\Delta t^{BA}_{equil.} . \qquad (19)$$

When the initial concentrations in each experiment are equal, $C^B_1/C^A_1 = 1$, as was the case here. A plot of $\ln\left(\dfrac{C^B_{1+N}}{C^A_{1+N}}\right)$ versus $-N\Delta t^{BA}_{equil.}$ yields a straight line with a slope of k = 1.0±0.4x10^{-3} s^{-1} (Figure 2f inset). Note that it is necessary to use the earlier scans (1-4 here, this depends on ligand conversion kinetics) in the experiment where the concentrations of ligand can be more accurately fit, as the relative error in determining ligand concentrations is high for later scans.

   Similarly, the rate constant can be calculated from a minimum number of 4 forward scans (Supplementary Scheme 1). A two-forward scan experiment with a short equilibration time at high temperature is performed first (Experiment A), and, in the case of thermolabile ligands like cocaine, C^A_1 can be assumed to be what was loaded into the calorimeter and the C^A_2 can be extracted from the second forward scan. An additional two-scan experiment (Experiment B) with a longer equilibration time at the highest temperature gives C^B_1 and C^B_2, from which the rate constant can be calculated using the $\Delta t^{BA}_{equil.}$ and Supplementary Eq. 19.
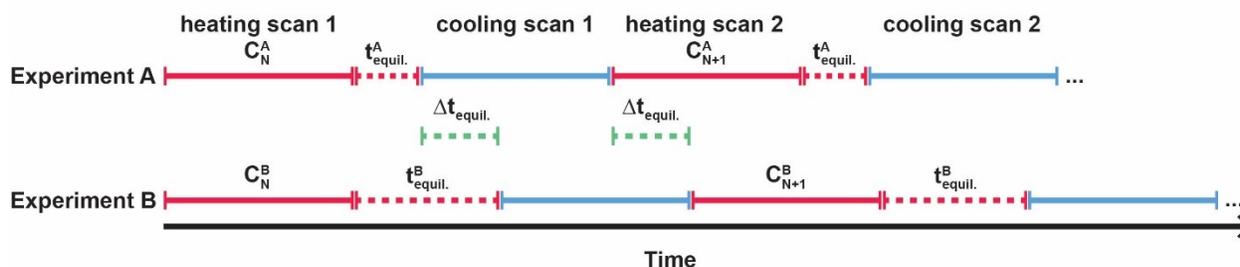
Supplementary Scheme 1. Time evolution of two DSC experiments with different high temperature equilibration periods. Heating and cooling scans are shown by solid red and blue increments respectively. Experiment A and B have shorter and longer high temperature equilibration times respectively, shown as dashed red increments. The difference in equilibration times is indicated by dashed green increments.
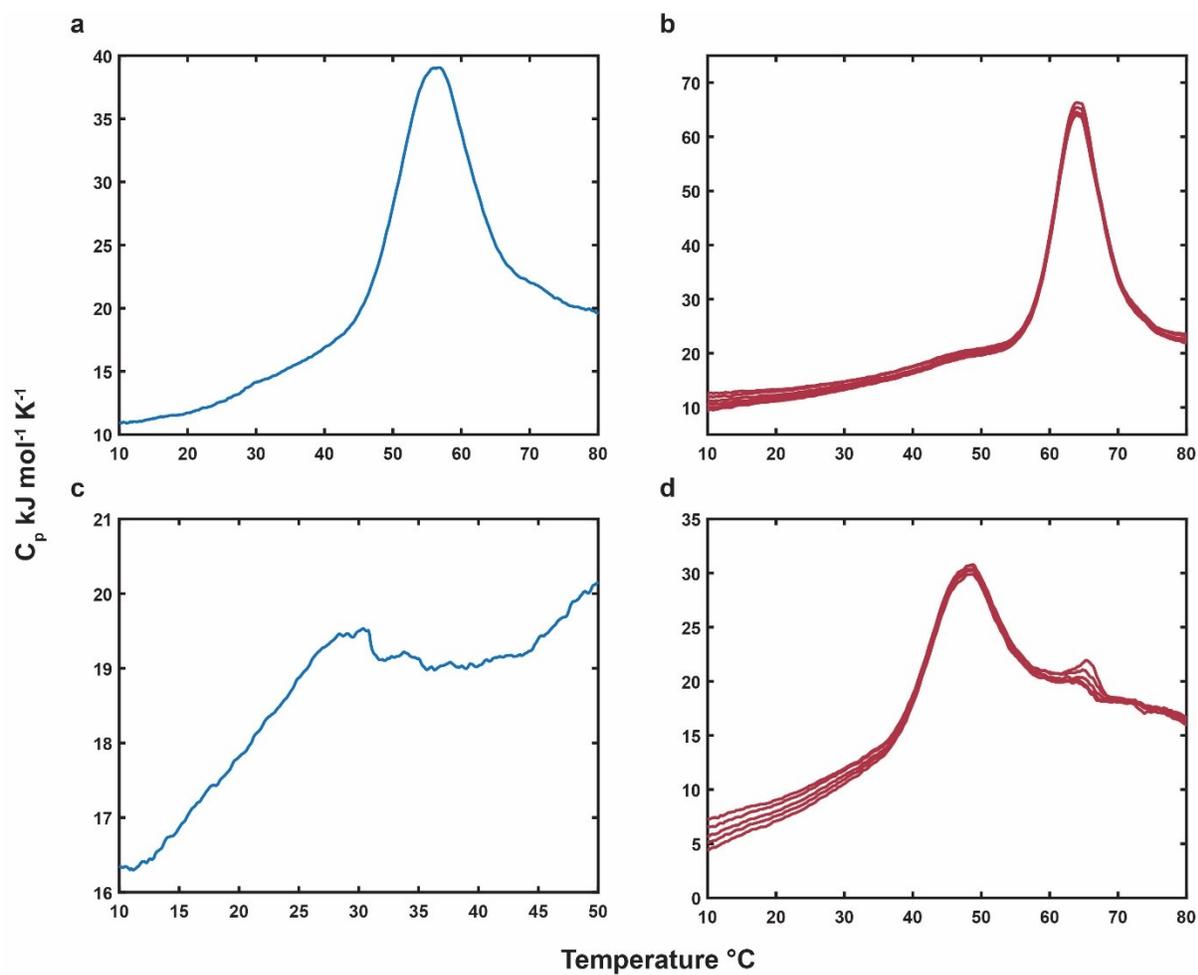
Supplementary Figure 1. Free and quinine-bound aptamers. (a) DSC profile for free MN4. (b) DSC profiles for MN4 bound to quinine. Successive scans show no change in profile. (c) DSC profile for free MN19. MN19 is largely unstructured in its free state. (d) DSC profiles for MN19 bound to quinine. Successive scans show no change in profile.
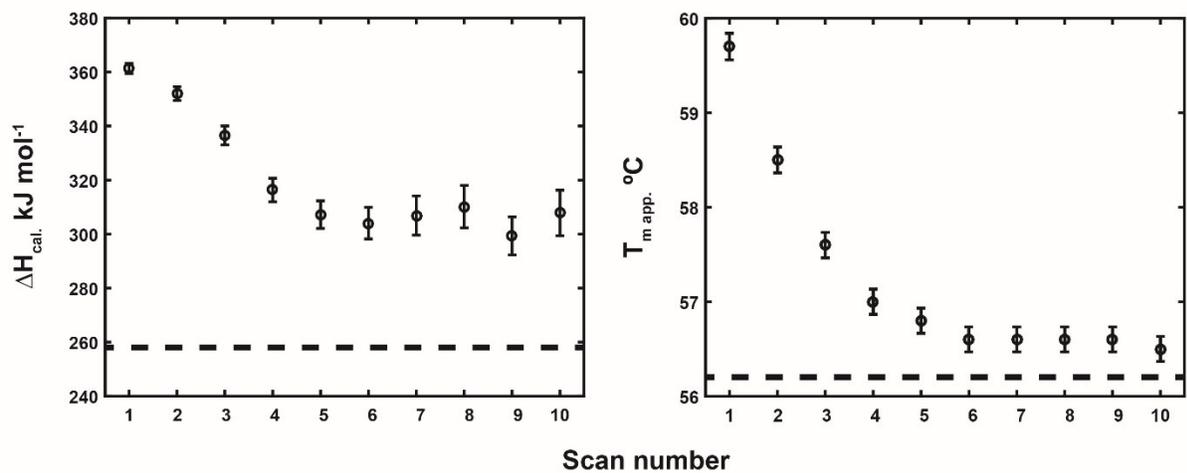
Supplementary Figure 2. Evidence for benzoylecgonine binding. Calorimetric enthalpy and apparent melting temperature versus forward scan number for the cocaine-bound MN4 dataset. Dashed black lines indicate the calorimetric enthalpy and T_m for free MN4 respectively.
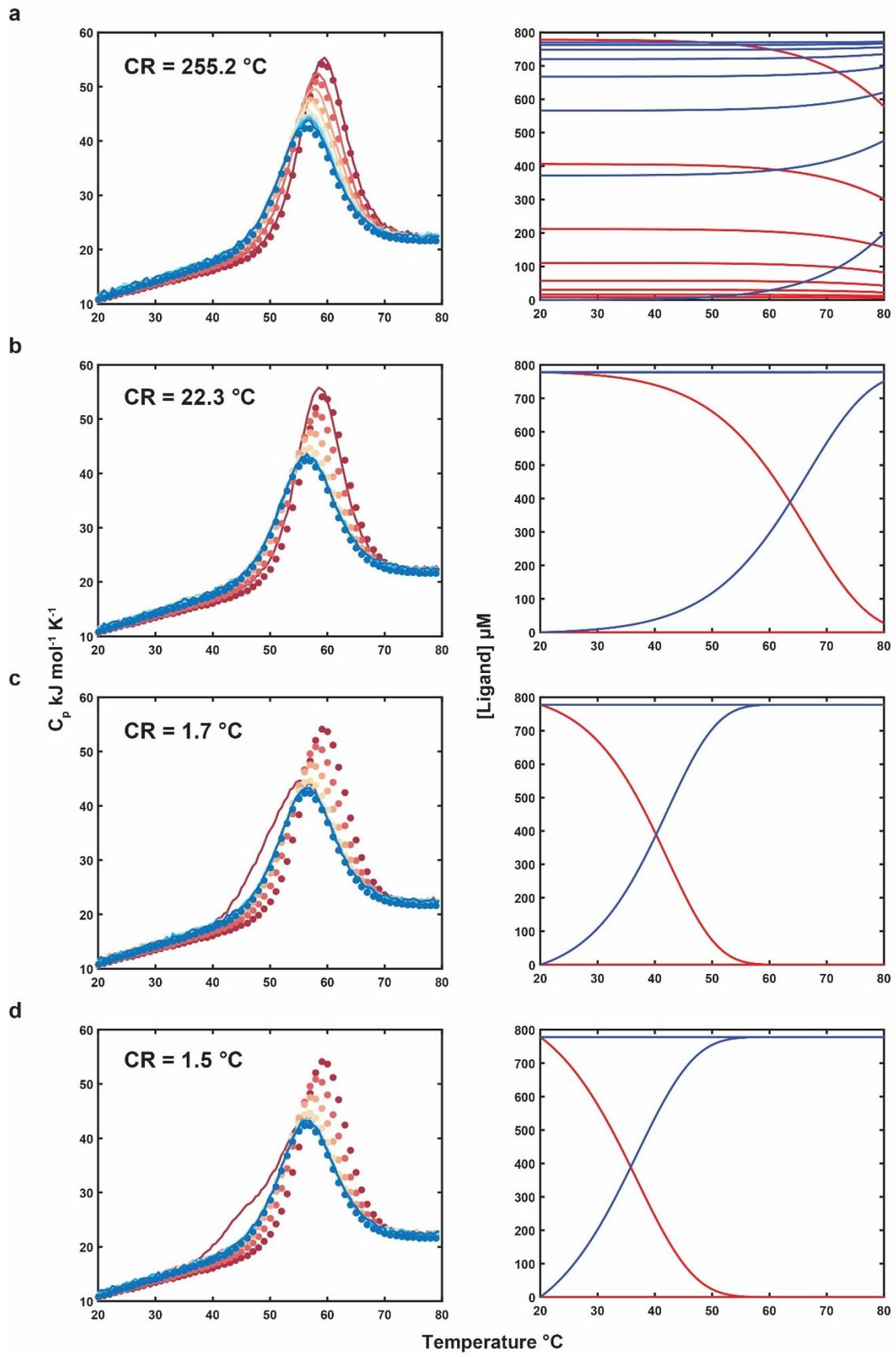
Supplementary Figure 3. Effects of scan rate and continuously-varying ligand conversion kinetics on thermolabile ligand DSC profiles. (a) Simulated DSC profiles with 1 °C min^{-1} scan rate and slow ligand conversion at low temperatures (E_a = 95.9 kJ mol^{-1}). (b) Simulated DSC profiles with 1 °C min^{-1} scan rate and moderate ligand conversion at low temperatures (E_a = 89.0 kJ mol^{-1}). (c) 0.005 °C min^{-1} scan rate and slow ligand conversion at low temperatures (E_a = 95.9 kJ mol^{-1}). (d) 1 °C min^{-1} scan rate and rapid ligand conversion at low temperatures (E_a = 81.0 kJ mol^{-1}). In (a-d), simulated DSC profiles where the concentration of cocaine is fixed (i.e. using the optimized parameters from fits of experimental data to our model) through each transition are shown as colored circles while data simulated with continuously-varying ligand concentrations are shown as solid curves. First and last scans are shown in dark red and dark blue. Gaussian noise was added to the simulated scans using the standard deviation of the high temperature experimental baseline. Concentrations of ligand 1 and 2 as a function of temperature are shown as red and blue lines respectively in the right-hand panels. First order kinetics in (a-d) were calculated with a pre-exponential factor estimated from literature rate constants (A = 7.51×10^{10} s^{-1})[1]. The conversion ratios (CR) given in the upper left corners of (a-d) were calculated as CR = scan rate/k(T$_m$$^{app.}$). Values of CR below ~20 °C give distortions of the first scan's shape and violate the assumption that most of the ligand conversion happens at high temperatures. Subsequent scans superimpose as the initial ligand is depleted entirely within the first scan.
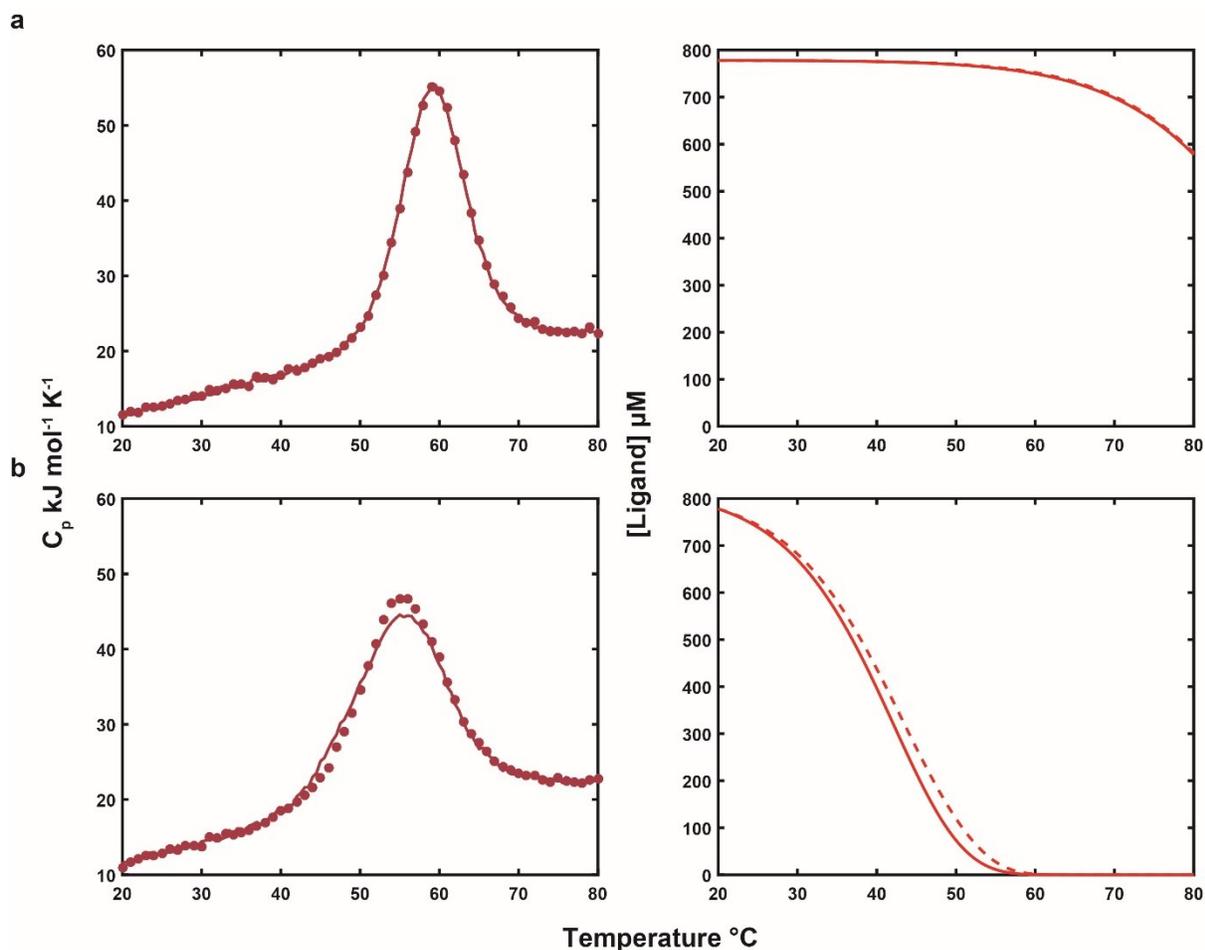
Supplementary Figure 4. Protection of the ligand by the biomolecule. (a) Simulated DSC profiles at 1 °C min⁻¹ scan rate for continuously-varying ligand conversion in the absence (dark red line) and presence (dark red circle) of protection of the ligand in the biomolecule binding pocket. (b) DSC profiles simulated at 0.005 °C min⁻¹ scan rate for continuously-varying ligand conversion in the absence (dark red line) and presence (dark red circle) of protection of the ligand in the biomolecule binding pocket. If the ligand can be protected by the biomolecule and the ligand is in excess, the rate constant for ligand conversion is given by $k^{app.}(T) = k(T) \times P_{free}(T)$. Ligand concentrations for the protected (dashed red line) and unprotected (red line) cases in (a) and (b) are shown in the panels immediately to the right. Rate constants for ligand conversion were calculated using E_a = 95.9 kJ mol⁻¹ and A = 7.51×10¹⁰ s⁻¹.
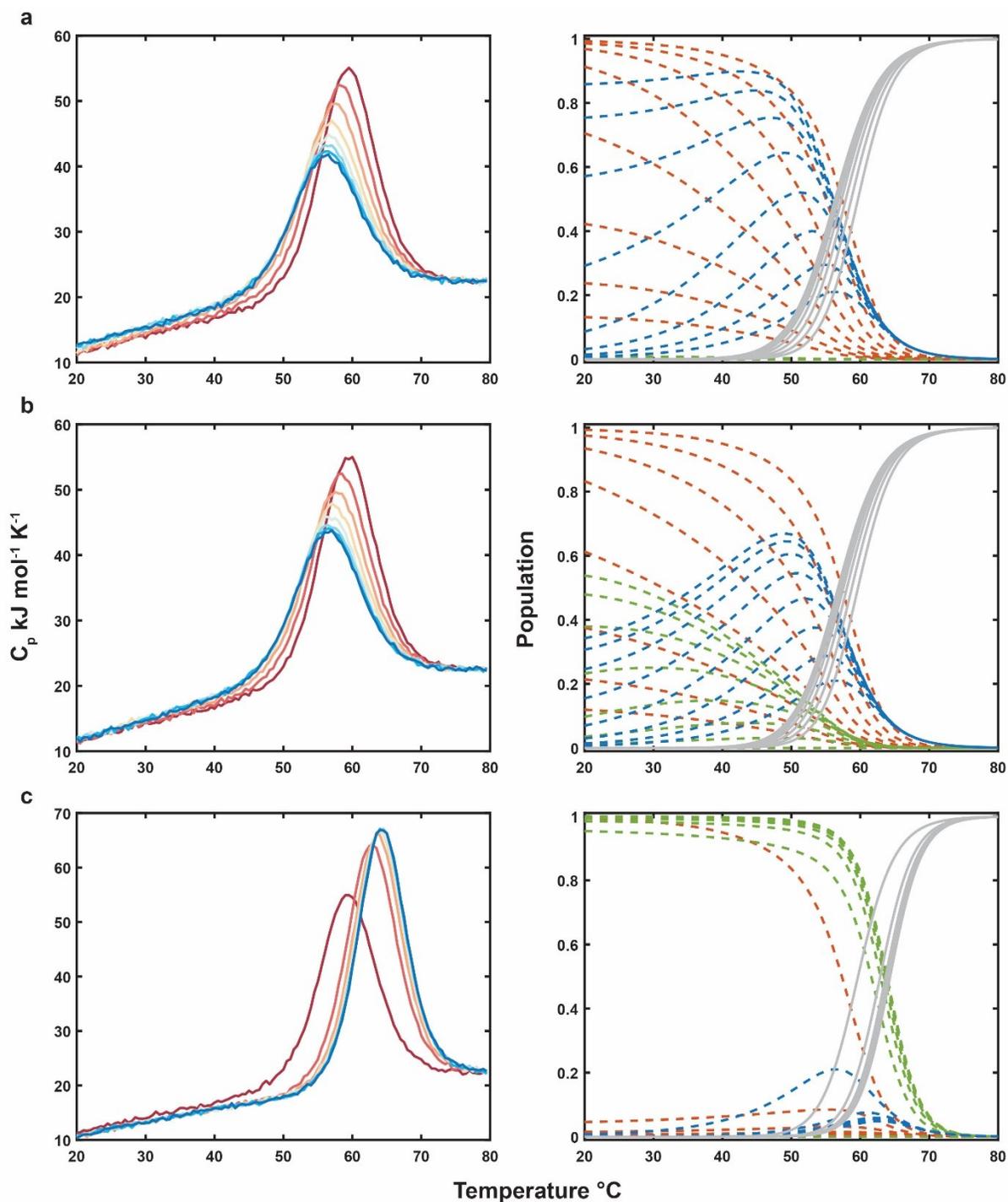
Supplementary Figure 5. Simulation of ligand binding scenarios. (a) Conversion of initial bound ligand to a form with negligible affinity for the aptamer (i.e. decrease in total bound ligand concentration). ΔH^{B1F} = -61.4 kJ mol^{-1}, ΔS^{B1F} = -108.1 J mol^{-1} K^{-1}. ΔH^{B2F} = -2.0 kJ mol^{-1}, ΔS^{B2F} = 16.0 J mol^{-1} K^{-1}. (b) Conversion of the initial ligand to a weaker binding form. ΔH^{B1F} = -61.4 kJ mol^{-1}, ΔS^{B1F} = -108.1 J mol^{-1} K^{-1}. ΔH^{B2F} = -14.0 kJ mol^{-1}, ΔS^{B2F} = 16.0 J mol^{-1} K^{-1}. (c) Conversion of the initial ligand to a tighter binding form. ΔH^{B1F} = -61.4 kJ mol^{-1}, ΔS^{B1F} = -108.1 J mol^{-1} K^{-1}. ΔH^{B2F} = -68.4 kJ mol^{-1}, ΔS^{B2F} = -103.0 J mol^{-1} K^{-1}. In each simulated profile, the [aptamer] = 83

µM, [ligand]$_{initial}$ = 778 µM, ΔH^{UF} = 271.3 kJ mol^{-1}, ΔS^{UF} = 824.2 J mol^{-1} K^{-1}, ΔC_p^{UF} = 0, ΔC_p^{B1F} = -1.5 kJ mol^{-1} K^{-1}, ΔC_p^{B2F} = -2.2 kJ mol^{-1} K^{-1}, and k$_{conversion}$ = 5×10^{-3} s^{-1}. Ligand was assumed to convert as a first order process at high temperature for 120 seconds. Heat capacity baselines were calculated as 12.8 + 0.292(T − T$_0$) - 0.0022(T − T$_0$)^2. Gaussian noise was added to the simulated profiles using the standard deviation of horizontal high temperature experimental baselines. Simulated DSC scans are shown as colored lines, where dark red and dark blue indicate first and last scans respectively. Populations for each ligand binding scenario are shown in the panels immediately to the right. The populations of initial ligand bound, converted ligand bound, and folded states are shown as orange, green, and blue dashed lines respectively. Populations of the unfolded state are shown as grey solid lines.

Supplementary Table 1. Cocaine concentrations extracted from global analysis of the cocaine-added MN4 datasets assuming benzoylecgonine can bind the aptamer.

Scan number	[Cocaine] µM, 120 second equilibrations	[Cocaine] µM, 600 second equilibrations
1	778.0±19.8	778.0±19.8
2	361.3±4.0	250.3±4.8
3	159.7±2.7	68.3±3.3
4	57.4±2.1	13.3±2.8
5	16.3±1.8	3.1±2.5
6	1.3±1.2	2.1±2.3
7	0.9±1.0	3.2±2.5

References

1.   J. B. Murray and H. I. Alshora, *J Clin Pharmacy*, 1978, **3**, 1-6.