Supplementary Information

Nearest-neighbor parameters for predicting DNA duplex stability in diverse molecular crowding conditions

Saptarshi Ghosh,[a] Shuntaro Takahashi,[a] Tatsuya Ohyama,[a] Tamaki Endoh,[a] Hisae Tateishi-Karimata,[a] and Naoki Sugimoto*[a][b]

[a]Frontier Institute for Biomolecular Engineering Research (FIBER), [b]Graduate School of Frontiers of Innovative Research in Science and Technology (FIRST), Konan University, Kobe, Japan

Contents

www.pnas.org/cgi/doi/10.1073/pnas.1920886117

Materials and methods

Materials. All DNA oligonucleotides designed for this study were purchased from Japan Bio Services Co., Ltd. (Saitama, Japan) and purified using high-performance liquid chromatography (HPLC). The DNA samples were dissolved in Milli-Q water and stored in -20 °C. Oligonucleotide concentrations were determined by measuring absorbance at 260 nm at 90 ºC using extinction coefficients. Polyethylene glycol 200 (PEG 200), ethylene glycol and 1,3-propanediol were purchased from Wako Pure Chemical Industries, Ltd. (Osaka, Japan) and used without further purification. Disodium hydrogen phosphate (Na_2HPO_4) and sodium chloride (NaCl) (Wako Pure Chemical Industries, Ltd.) and disodium ethylenediaminetetraacetate (Na_2EDTA) (Dojindo Molecular Technologies Inc., Kumamoto, Japan) were used as received.

Selection of DNA duplexes. All sequences used to determine the NN parameters exhibited a two-state transition. Very short DNA sequences could not be selected because the stability of DNA duplexes decreased under the crowded conditions of PEG 200 and the experiments proceeded at the physiologically relevant NaCl concentration of 100 mM. On the other hand, the NN model deviated for longer sequences due to non-two-state transitions (1). Therefore, we selected DNA sequences between 8 and 16 mer, thus ensuring sufficient stability and two-state transitions. The sequences had different combinations of NN frequencies, covering all 10 NN base pairs and two initiation factors for DNA duplex formation. The NN in the total set of designed oligonucleotides occurred at the following frequencies: d(AA/TT) = 26, d(AT/TA) = 30, d(TA/AT) = 19, d(CA/GT) = 44, d(GT/CA) = 44, d(CT/GA) = 38, d(GA/CT) = 48, d(CG/GC) = 43, d(GC/CG) = 30 and d(GG/CC) = 32.

Circular dichroism (CD) measurements. The CD spectra were measured using a J-1500 spectropolarimeter (Jasco Corporation, Hachioji, Japan) equipped with a temperature controller. All spectra were collected at 4 °C. The chamber holding the cuvette was constantly flushed with a stream of dry N_2 gas to avoid water condensation on the cuvette

exterior. Spectra were measured from 200 to 340 nm in 0.1 cm path-length cuvettes, at a scan rate of 50 nm min^{-1}. The concentrations of the samples were maintained at 20 μM in a buffer containing 100 mM NaCl, 10 mM Na_2HPO_4 (pH 7.0) and 1 mM Na_2EDTA with or without 40 wt% PEG 200.

UV melting studies. Absorption spectra were obtained on a Shimadzu 1800 spectrophotometer equipped with a thermoprogrammer. All experiments proceeded in a buffer containing 100 mM NaCl, 10 mM Na_2HPO_4 and 1 mM Na_2EDTA in the presence of 40 wt% PEG 200. We adjusted the pH of the buffer to 7.0 after adding the cosolutes. Oligonucleotide concentrations were varied over a 50 – 100-fold range for melting experiments. Sample solutions were maintained at 9ºC for 5 min, followed by a decrease in temperature to 0 ºC at a rate of 1 ºC min^{-1} to anneal the duplexes. The samples were maintained at 0 ºC for 5 min, then heated to 90 ºC at a rate of 0.5 ºC min^{-1} to melt the duplexes. The absence of hysteresis between the denaturation and renaturation profiles indicated that the transition between the duplex and single-strand was two-state. Water condensation on the cuvette exterior at low temperature was avoided by flushing with a constant stream of dry N_2 gas.

Thermodynamic analysis. Thermodynamic parameters ($\Delta H°$, $\Delta S°$ and $\Delta G°_{37}$) for DNA duplexes were determined from the T_m^{-1} vs. ln (C_t/s) plots, as we previously described (2). Thermodynamic parameters were calculated from the slope and intercept of the linear plots using the following equations:

$$T_m^{-1} = R \ln (C_t/s)/\Delta H° + \Delta S°/\Delta H° \qquad \text{(S1)},$$

$$\Delta G°_{37} = \Delta H° - 310.15 \bullet \Delta S° \qquad \text{(S2)},$$

where, R is the gas constant, C_t is the total strand concentration of the oligonucleotides and s reflects the sequence symmetry of the self (s = 1) or non-self-complementary strands (s = 4). Following the standard practice for calculation of the thermodynamic parameters, we assumed that the difference in heat capacity (ΔC_p) between the two states (single-strand and duplex) was zero (3-7). Because the T_m for most of the studied sequences under

crowded conditions were not far from 37 ºC, zero ΔC_p approximation should be acceptable for $\Delta G°_{37}$ calculations due to minimal extrapolations. Generally, $\Delta G°_{37}$ is relatively insensitive to ΔC_p due to enthalpy-entropy compensation (8-10).

Calculation of nearest-neighbor parameters

According to the nearest-neighbor (NN) model, the free energy change of duplex formation consisted of three terms. (i) Free energy change for helix propagation as the sum of each subsequent base pair. (ii) Free energy change for helix initiation to form a first base pair in the double helix. Since helix initiation can occur either by G•C or A•T pairing, two initiation factors are considered. These two types of terminal pairs contributed differently to the total free energy change for DNA duplexes (10). (iii) Free energy changes due to entropic penalty for the maintenance of the C_2 symmetry for self-complementary sequences. The 10 NN base pairs and two initiation factors were determined under crowding conditions using thermodynamic data for DNA duplexes in the presence of PEG 200 obtained from UV melting experiments and linear least squares software written in Python. The program calculated ΔG^o_{37}, ΔS^o and ΔH^o using a set of 13 parameters (ten Watson-Crick NN base pairs, two terminal pairs and a symmetry parameter for self-complementary sequences). In the algorithm used to determine the parameters, the initial base parameters (P^0) without symmetry correction parameter were set to zero. The i was group index for the parameter set.

$$P^0 = \{p_1, p_2, \cdots, p_{13}\} = \{0.0, 0.0, \cdots, 0.0\} \quad (S3)$$

The parameters for symmetry correction were set to the same value as that in the absence of cosolutes, since it does not depend on the environment, but on whether the sequence is self- or non-self-complementary. Modified parameter sets, such as positive ($P_{i,1}$) and negative ($P_{i,-1}$) parameter sets, were prepared based on $P_{i,0}$. $P_{i,0}$ has the same value as P^0. The $P_{i,1}$ and $P_{i,-1}$ had larger or smaller parameters, respectively, determined by specific small value (ΔE) compared with $P_{i,0}$. The ΔE was set to 0.1.

$$
\begin{aligned}
P_1 &= \begin{cases} P_{1,0} & = \{p_1, p_2, \cdots, p_{13}\} = \{0.0, 0.0, \cdots, 0.0\} \\ P_{1,1} & = \{p_1 + \Delta E, p_2, \cdots, p_{13}\} = \{0.1, 0.0, \cdots, 0.0\} \\ P_{1,-1} & = \{p_1 - \Delta E, p_2, \cdots, p_{13}\} = \{-0.1, 0.0, \cdots, 0.0\} \end{cases} \\
P_2 &= \begin{cases} P_{2,0} & = \{p_1, p_2, \cdots, p_{13}\} = \{0.0, 0.0, \cdots, 0.0\} \\ P_{2,1} & = \{p_1, p_2 + \Delta E, \cdots, p_{13} = \{0.0, 0.1, \cdots, 0.0\}\} \\ P_{2,-1} & = \{p_1, p_2 - \Delta E, \cdots, p_{13}\} = \{0.0, -0.1, \cdots, 0.0\} \end{cases} \\
&\vdots \\
P_{13} &= \begin{cases} P_{13,0} & = \{p_1, p_2, \cdots, p_{13}\} = \{0.0, 0.0, \cdots, 0.0\} \\ P_{13,1} & = \{p_1, p_2, \cdots, p_{13} + \Delta E\} = \{0.0, 0.0, \cdots, 0.1\} \\ P_{13,-1} & = \{p_1, p_2, \cdots, p_{13} - \Delta E\} = \{0.0, 0.0, \cdots, -0.1\} \end{cases}
\end{aligned} \tag{S4}
$$

Since $P_{i,j}$ $(j = -1, 0, 1)$ was prepared for each parameter, the total number of parameter sets were 39. The ΔG^o_{37} value for each sequence was calculated using these parameter sets and compared with error sums of ΔG^o between the experimental and predicted, as per the following equation:

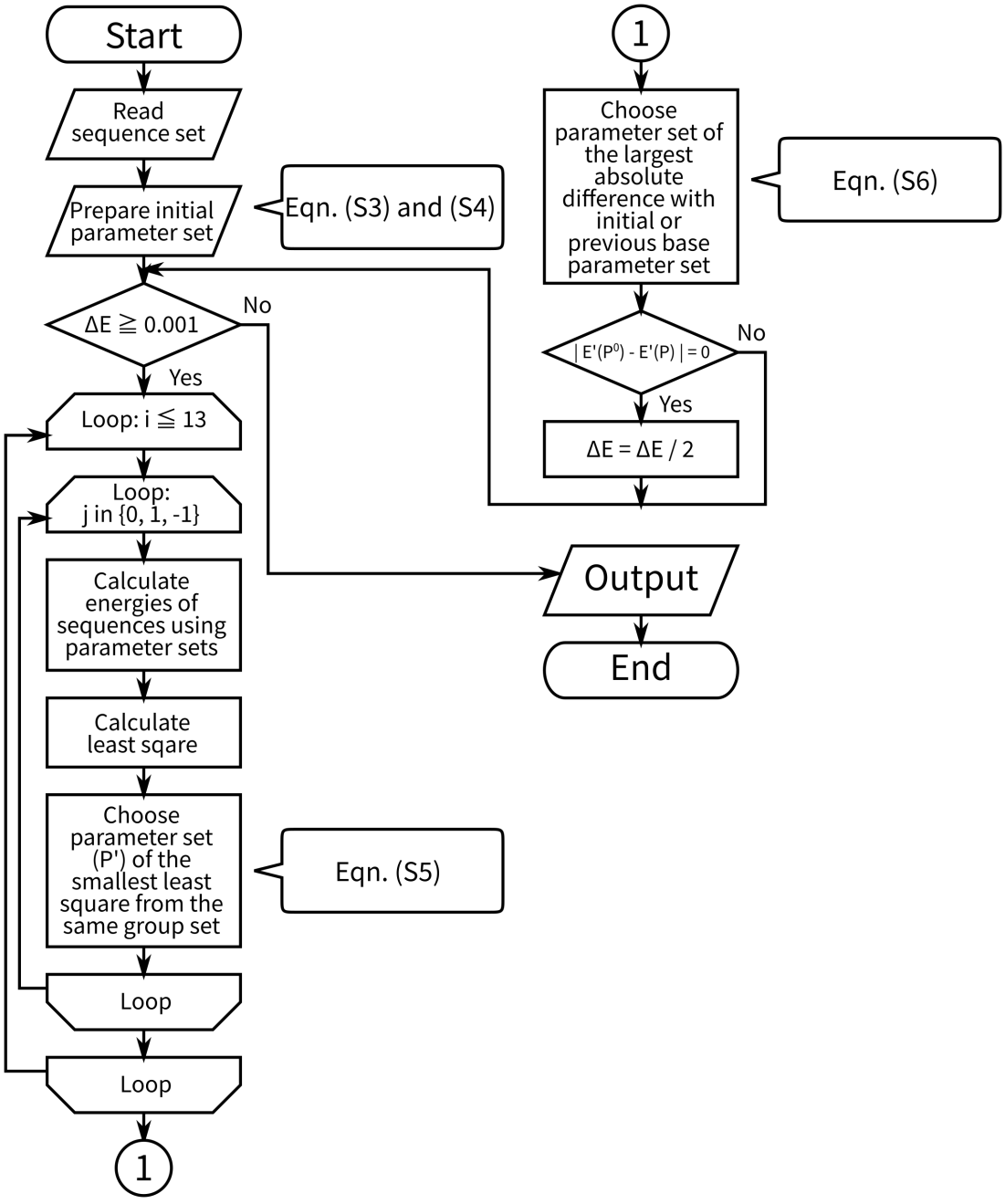
$$
P_i = \underset{P \in \{P_{i,0}, P_{i,1}, P_{i,-1}\}}{\operatorname{argmin}} \sum_k \left(E_k - E_k^{'}(P) \right)^2 \tag{S5}
$$

where, k, and P_i denote each sequence and representative parameter sets, respectively. E_i and $E'_i(P)$ are the experimental and predicted values, respectively. The parameter set minimizing $\sum_k \left(E_k - E'_k(P) \right)^2$ was adopted as the representative parameter set (P_i). Among the representative parameter sets, P_i having a large absolute difference of the $E'(P_i)$ with prediction energy ($E'(P^0)$) using initial or previous base parameter set (P^0) was a parameter set that has greatly contributed to the optimization. The parameter set was adopted as the next base parameter (P') as follows:

$$
P' = \underset{P \in P_i}{\operatorname{argmax}} \left| E'(P^0) - E'(P) \right| \tag{S6}
$$

Based on P', new positive and negative parameter sets were prepared and used to calculate ΔG^o. These calculations were repeated until $|E'(P^0) - E'(P)|$ remained unchanged. When the absolute difference did not change, positive and negative parameter sets used half of ΔE $(\Delta E/2)$. These procedures were repeated until the ΔE was < 0.001. The method of calculating ΔH^o was the same as that described above. The ΔS^o values were

obtained from the determined values of ΔG^o_{37} and ΔH^o using equation S2. For better comprehension, we provide a schematic representation of the abovementioned procedure below.



Scheme S1. Calculation procedure of nearest-neighbor parameter

Water activity measurement

Water activity was determined by vapor phase osmometry using a pressure osmometer (Wescor Inc., South Logan, UT, USA) or by freezing point depression osmometry using a Knauer osmometer **(**Knauer Wissenschaftliche Geräte GmbH, Berlin, Germany).

Water activity (a_w) was calculated from the measured osmolality (mmol kg^{-1}) using the following equation (11):

$$\Psi = (RT/M_w) \ln a_w \qquad \text{(S7)},$$

where Ψ and M_w represent the water potential and the molecular weight of water (0.018 kg mol^{-1}), respectively. The relationship between water potential and osmolality is given by the equation (11):

$$\Psi \text{ (MPa)} = \text{osmolality (mmol kg}^{-1}\text{)} \times 10^3 / (-400) \qquad \text{(S8)}.$$

Cell lysate experiment

For melting assays in cell lysate, HPLC purified fluorescence-labeled DNA oligonucleotides were purchased from Japan Bio Services Co., Ltd. (Saitama, Japan). 5´ end of the DNA was labelled with 6FAM (5´-6FAM-AGTAACGCCAT-3´) and 3´ end of the complementary sequence with quencher BHQ1 (5´-ATGGCGTTACT-BHQ1-3´). We used lysates from HeLa cells (Santa Cruz Biotechnology, USA) containing nuclear extract (250 µg in 50 µl) in 20 mM HEPES buffer (pH 7.9), 20% v/v glycerol, 100 mM KCl, 0.2 mM EDTA, 0.5 mM phenyl methyl sulfonyl fluoride (PMSF) and 0.5 mM dithiothreitol (DTT). To avoid the effect of glycerol present in the lysate solution, we dialyzed the lysate solution overnight using the Slide-A-Lyzer (MWCO 2000) dialysis kit against the same buffer containing all the components of the lysate buffer except the glycerol. To prepare a 5× concentrated lysate, we first dialyzed the lysate against 4 mM HEPES buffer (pH 7.9), 20 mM KCl, 0.04 mM EDTA, 0.1 mM PMSF and 0.1 mM DTT. The lysate was then concentrated to one-fifth of the original volume, such that all the buffer composition remained same as that of the original lysate solution. We concentrated the lysate by gradual evaporation (12). Samples were evaporated in a centrifugal vacuum evaporator at 0.1 MPa pressure. The volume was repeatedly monitored

with time and the final volume was measured using a well-calibrated micropipette. We collected the fluorescence data for the fluorescence-labeled DNA using Mx Pro 300 5P qPCR. Thermodynamic parameters were obtained by fitting the melting curve in a two-state model (3-5).

Table S1. Non-self-complementary DNA sequences with nearest-neighbor frequencies.

		Nearest-neighbor combinations present in duplex									
No.	Sequence[a]	dAA dTT	dAT dTA	dTA dAT	dCA dGT	dGT dCA	dCT dGA	dGA dCT	dCG dGC	dGC dCG	dGG dCC
NS1	d(GGCAGTTC)	1	0	0	1	1	1	1	0	1	1
NS2	d(GGTTCAGC)	1	0	0	1	1	1	1	0	1	1
NS3	d(CGCTGTAG)	0	0	1	1	1	2	0	1	1	0
NS4	d(CGTGCTAG)	0	0	1	1	1	2	0	1	1	0
NS5	d(AGTAACGCCAT)	1	1	1	1	2	1	0	1	1	1
NS6	d(AATGCCGTAGT)	1	1	1	1	2	1	0	1	1	1
NS7	d(CCATCGCTACC)	0	1	1	1	1	1	1	1	1	2
NS8	d(CGATGGCCTAC)	0	1	1	1	1	1	1	1	1	2
NS9	d(CGCTTGTTAC)	2	0	1	1	2	1	0	1	1	0
NS10	d(CCGTAACGTTGG)	2	0	1	1	3	0	0	2	0	2
NS11	d(ACTGACTGACTG)	0	0	0	3	3	3	2	0	0	0
NS12	d(ACTGACTGACTGACTG)	0	0	0	4	4	4	3	0	0	0

[a]Sequences NS1 and NS2, NS3 and NS4, NS5 and NS6 and NS7 and NS8 have identical NN sets.

Table S2. Reproduction of the published NN parameters of DNA duplex formation in a solution without cosolute by our software[a].

	Parameters calculated using our software			Our previously reported parameters (4)		
Sequence	$\Delta H°$ (kcal mol^{-1})	$\Delta S°$ (cal mol^{-1}K^{-1})	$\Delta G°_{37}$ (kcal mol^{-1})	$\Delta H°$ (kcal mol^{-1})	$\Delta S°$ (cal mol^{-1}K^{-1})	$\Delta G°_{37}$ (kcal mol^{-1})
d(AA/TT)	–8.0	–21.9	–1.2	–8.0	–21.9	–1.2
d(AT/TA)	–5.6	–15.2	–0.9	–5.6	–15.2	–0.9
d(TA/AT)	–6.6	–18.4	–0.9	–6.6	–18.4	–0.9
d(CA/GT)	–8.2	–21.0	–1.7	–8.2	–21.0	–1.7
d(GT/CA)	–6.7	–16.8	–1.5	–6.6	–16.4	–1.5
d(CT/GA)	–8.7	–23.1	–1.5	–8.8	–23.5	–1.5
d(GA/CT)	–9.4	–25.5	–1.5	–9.4	–25.5	–1.5
d(CG/GC)	–11.8	–29.0	–2.8	–11.8	–29.0	–2.8
d(GC/CG)	–10.4	–26.0	–2.4	–10.5	–26.4	–2.3
d(GG/CC)	–10.7	–27.8	–2.1	–10.9	–28.4	–2.1
Initiation	0.6	–9.0	3.4	0.6	–9.0	3.4
Symmetry	0	–1.4	0.4	0	–1.4	0.4

[a]We used thermodynamic data for DNA sequences in a buffer containing 1 M NaCl without cosolute as we described (4) to obtain reported values.

Table S3. Thermodynamic parameters for DNA helix initiation and propagation in 100 mM NaCl without cosolute[a].

Sequence	$\Delta H°$ (kcal mol^{-1})	$\Delta S°$ (cal mol^{-1} K^{-1})	$\Delta G°_{37}$ (kcal mol^{-1})
d(AA/TT)	–7.9	–23.3	–0.65
d(AT/TA)	–7.2	–21.3	–0.60
d(TA/AT)	–7.2	–22.0	–0.36
d(CA/GT)	–8.5	–23.4	–1.23
d(GT/CA)	–8.4	–23.2	–1.20
d(CT/GA)	–7.8	–21.5	–1.11
d(GA/CT)	–8.2	–23.4	–0.93
d(CG/GC)	–10.6	–28.2	–1.85
d(GC/CG)	–9.8	–25.0	–2.05
d(GG/CC)	–8.0	–20.4	–1.69
Initiation per GC	0.1	–2.8	0.98
Initiation per AT	2.3	4.1	1.03
Self-complementary	0	–1.4	0.40
Non-self-complementary	0	0	0

[a]Values are corrected for 100 mM NaCl from 1 M NaCl values as described by Huguet *et al*. (13). NN parameters in 1 M NaCl are as described by SantaLucia *et al*. (10).

Table S4. Measured and predicted thermodynamic parameters for DNA duplex formation with 40 wt% PEG 200 and 100 mM NaCl in 10 mM phosphate buffer (pH 7.0).

No	Sequence[a]	Measured				Predicted			
		$\Delta H°$ (kcal mol^{-1})	$\Delta S°$ (cal mol^{-1}K^{-1})	$\Delta G°_{37}$ (kcal mol^{-1})	T_m[b] (ºC)	$\Delta H°$ (kcal mol^{-1})	$\Delta S°$ (cal mol^{-1}K^{-1})	$\Delta G°_{37}$ (kcal mol^{-1})	T_m[b] (ºC)
NS1	d(GGCAGTTC)	–67.4	–200.9	–5.1	30.8	–65.7	–195.0	–5.2	30.9
NS2	d(GGTTCAGC)	–61.0	–180.2	–5.1	29.4	–65.7	–195.0	–5.2	30.9
NS3	d(CGCTGTAG)	–64.3	–189.9	–5.4	31.7	–64.3	–191.4	–5.0	29.5
NS4	d(CGTGCTAG)	–68.1	–202.8	–5.2	31.9	–64.3	–191.4	–5.0	29.5
NS5	d(AGTAACGCCAT)	–83.3	–246.3	–6.9	38.3	–75.5	–222.5	–6.5	36.8
NS6	d(AATGCCGTAGT)	–77.8	–229.2	–6.7	37.8	–75.5	–222.5	–6.5	36.8
NS7	d(CCATCGCTACC)	–93.5	–274.4	–8.4	43.4	–83.1	–242.6	–7.8	42.0
NS8	d(CGATGGCCTAC)	–95.8	–280.8	–8.7	44.1	–83.1	–242.6	–7.8	42.0
NS9	d(CGCTTGTTAC)	–76.4	–226.3	–6.2	35.2	–83.2	–249.9	–6.0	35.1
NS10	d(CCGTAACGTTGG)	–95.0	–281.5	–8.7	44.2	–99.2	–292.0	–8.6	43.7
NS11	d(ACTGACTGACTG)	–88.4	–259.2	–8.0	42.2	–92.6	–272.9	–8.0	41.9
NS12	d(ACTGACTGACTGACTG)	–124.9	–364.3	–11.9	51.0	–123.2	–359.6	–11.6	50.5
S1	d(GGACGTCC)	–63.0	–186.0	–5.3	35.4	–63.0	–185.7	–5.4	35.7
S2	d(GACCGGTC)	–64.1	–189.3	–5.4	35.6	–63.0	–185.7	–5.4	35.7
S3	d(CGTCGACG)	–63.9	–186.4	–6.1	39.0	–67.9	–200.6	–5.7	37.0
S4	d(CGACGTCG)	–63.2	–183.8	–6.2	39.3	–67.9	–200.6	–5.7	37.0
S5	d(CAAGCTTG)	–79.7	–245.4	–3.6	28.9	–70.5	–213.8	–4.2	30.6
S6	d(CTTGCAAG)	–72.5	–221.2	–3.9	29.6	–70.5	–213.8	–4.2	30.6
S7	d(CGGTACCG)	–69.5	–209.3	–4.6	32.3	–63.9	–189.1	–5.2	35.0
S8	d(CCGTACGG)	–60.5	–177.7	–5.4	34.7	–63.9	–189.1	–5.2	35.0
S9	d(GATCCGGATC)	–82.5	–247.6	–5.7	37.3	–73.1	–216.8	–5.9	37.8
S10	d(GGATCGATCC)	–77.9	–231.8	–6.0	38.2	–73.1	–216.8	–5.9	37.8
S11	d(ATGAGCTCAT)	–71.8	–215.4	–5.0	34.3	–71.6	–215.1	–4.9	33.6
S12	d(ATCAGCTGAT)	–77.1	–232.8	–4.9	34.0	–71.6	–215.1	–4.9	33.6
S13	d(CATAGGCCTATG)	–86.0	–256.3	–6.5	39.8	–92.8	–278.3	–6.5	39.7
S14	d(CTATGGCCATAG)	–91.9	–274.4	–6.8	40.5	–92.8	–278.3	–6.5	39.7
S15	d(GCGAATTCGC)	–66.9	–193.1	–7.0	43.1	–73.9	–217.4	–6.5	40.4
S16	d(AGTCATGACT)	–69.8	–210.5	–4.5	32.3	–76.4	–230.0	–5.0	34.5
S17	d(GACGACGTCGTC)	–95.0	–277.0	–9.1	48.2	–96.1	–280.4	–9.2	48.6
S18	d(ATCGCTAGCGAT)	–59.8	–171.2	–6.7	43.1	–66.7	–192.2	–7.1	43.7
S19	d(GCAAGCCGGCTTGC)	–96.2	–270.8	–12.2	58.5	–97.4	–275.5	–12.0	58.4
S20	d(CGATCGGCCGATCG)	–90.9	–256.0	–11.5	56.9	–96.5	–274.6	–11.4	56.3

S21	d(CATATGGCCATATG)	–129.4	–394.0	–7.2	40.7	–131.1	–399.3	–7.2	40.8
S22	d(CAAGATCGATCTTG)	–116.4	–348.9	–8.2	44.6	–111.1	–331.5	–8.3	44.5
S23	d(CGCGTACGCGTACGCG)	–126.9	–365.3	–13.6	57.9	–116.9	–331.6	–14.1	60.9
S24	d(CGCAAGCCGGCTTGCG)	–109.4	–305.0	–14.8	64.2	–110.2	–307.8	–14.8	64.8

[a]DNA duplex consists of denoted DNA strand and its complementary DNA strand. [b]Melting temperatures were calculated for total strand concentration of 100 μM.

Table S5. Thermodynamic parameters of d(ATGCGCAT) in different cosolute solutions and their water activity values[a].

Solution	$\Delta H°$ (kcal mol^{-1})	$\Delta S°$ (cal mol^{-1} K^{-1})	$\Delta G°_{37}$ (kcal mol^{-1})	Water activity[b]
d(ATGCGCAT) in 1 M NaCl				
No cosolute	–62.2 ± 1.9	–171 ± 5.4	–9.2 ± 0.4	0.967
20 wt% ethylene glycol	–50.9 ± 2.4	–140 ± 7.9	–7.5 ± 0.3	0.907
30 wt% ethylene glycol	–47.3 ± 1.8	–131 ± 5.8	–6.7 ± 0.2	0.861
40 wt% ethylene glycol	–49.6 ± 1.9	–141 ± 7.9	–5.9 ± 0.4	0.832
5 wt% glycerol	–59.6 ± 2.2	–163 ± 7.0	–9.0 ± 0.3	0.954
10 wt% glycerol	–58.5 ± 1.4	–162 ± 4.2	–8.3 ± 0.4	0.938
15 wt% glycerol	–53.9 ± 1.2	–147 ± 3.9	–8.3 ± 0.3	0.930
20 wt% glycerol	–53.9 ± 1.3	–148 ± 3.1	–8.0 ± 0.4	0.917
25 wt% glycerol	–50.4 ± 1.1	–138 ± 3.7	–7.6 ± 0.3	0.899
30 wt% glycerol	–48.3 ± 1.7	–132 ± 4.8	–7.4 ± 0.3	0.891
40 wt% glycerol	–49.4 ± 1.5	–137 ± 5.0	–6.9 ± 0.3	0.866
50 wt% glycerol	–50.8 ± 2.2	–143 ± 4.2	–6.4 ± 0.4	0.843
15 wt% 2-methoxyethanol	–55.5 ± 1.2	–156 ± 3.5	–7.1 ± 0.4	0.927
20 wt% 2-methoxyethanol	–53.2 ± 1.3	–150 ± 3.7	–6.7 ± 0.3	0.915
40 wt% 2-methoxyethanol	–47.9 ± 0.9	–138 ± 3.7	–4.7 ± 0.3	0.875
15 wt% 1,3-propanediol	–62.8 ± 1.1	–176 ± 5.0	–8.0 ± 0.5	0.933
20 wt% 1,3-propanediol	–48.3 ± 1.2	–133 ± 3.6	–7.1 ± 0.5	0.918
40 wt% 1,3-propanediol	–48.8 ± 2.1	–138 ± 6.2	–6.0 ± 0.4	0.881
20 wt% 1,2-demethoxyethane	–49.9 ± 2.0	–138 ± 5.7	–7.1 ± 0.4	0.934
30 wt% 1,2-demethoxyethane	–49.1 ± 2.4	–138 ± 5.7	–6.3 ± 0.6	0.916
10 wt% PEG 200	–53.7 ± 1.7	–147 ± 5.6	–8.1 ± 0.2	0.953
20 wt% PEG 200	–51.4 ± 1.6	–143 ± 5.1	–7.0 ± 0.3	0.939
30 wt% PEG 200	–50.3 ± 1.8	–143 ± 5.6	–5.9 ± 0.2	0.923
40 wt% PEG 200	–47.8 ± 1.8	–138 ± 6.4	–5.0 ± 0.3	0.911
50 wt% PEG 200	–43.9 ± 2.3	–127 ± 8.2	–4.5 ± 0.4	0.896
20 wt% PEG 2000	–54.3 ± 2.7	–148 ± 8.7	–8.4 ± 0.4	0.956
40 wt% PEG 2000	–50.3 ± 1.5	–138 ± 5.8	–7.5 ± 0.4	0.948
10 wt% PEG 8000	–58.5 ± 3.2	–161 ± 10.0	–8.6 ± 0.4	0.959
15 wt% PEG 8000	–57.1 ± 2.7	–158 ± 8.4	–8.1 ± 0.5	0.955
d(ATGCGCAT) in 100 mM NaCl				
No cosolute	–58.8 ± 3.2	–161 ± 9.9	–8.9 ± 0.6	0.996
10 wt% PEG 200	–60.2 ± 3.5	–169 ± 2.5	–7.8 ± 0.2	0.985
20 wt% PEG 200	–56.4 ± 1.7	–160 ± 6.0	–6.8 ± 0.3	0.965
30 wt% PEG 200	–56.6 ± 3.3	–163 ± 9.0	–6.0 ± 0.3	0.956
40 wt% PEG 200	–51.2 ± 2.7	–148 ± 7.9	–5.3 ± 0.5	0.947

[a]Data were collected from our previously published report (14). [b]Water activities were calculated at 37 ºC from corresponding osmolality values of respective solutions calculated as described above.

Table S6. Osmolality and water activities in absence and presence of different cosolutes in 100 mM NaCl[a].

Cosolute	Osmolality (mmol kg^{-1})	Water activity
None	212 ± 1	0.996
20 wt% EG	3275 ± 2	0.943
20 wt% 1,3 PDO	2768 ± 2	0.952

[a]Experiments proceeded in a buffer containing 10 mM Na_2HPO_4 (pH 7.0), 1 mM Na_2EDTA, 100 mM NaCl and cosolute (as indicated). Data for 1,3 PDO and EG were obtained by freezing point depression osmometry and vapor phase osmometry, respectively.

Table S7. Nearest-neighbor parameters for $\Delta H^{\circ}_{[\text{cation}]}$ and $\Delta H^{\circ}_{[40\text{ wt\% PEG 200}]}$ in 100 mM NaCl along with prefactors (m_{cs}) for different cosolutes[a].

Sequence	$\Delta H^{\circ}_{[\text{cation}]}$ (kcal mol^{-1})	$\Delta H^{\circ}_{[40\text{wt\% PEG 200}]}$ (kcal mol^{-1})	$m_{\text{PEG/1,2 DME}}$ (kcal mol^{-1})	$m_{\text{EG/GOL}}$ (kcal mol^{-1})	$m_{\text{1,3 PDO/2-ME}}$ (kcal mol^{-1})
d(AA/TT)	–7.9	1.4	28	15	16
d(AT/TA)	–7.2	–2.2	–44	–24	–26
d(TA/AT)	–7.2	2.9	58	31	34
d(CA/GT)	–8.5	–4.6	–92	–50	–53
d(GT/CA)	–8.4	–0.8	–16	–9	–9
d(CT/GA)	–7.8	4.4	88	48	51
d(GA/CT)	–8.2	3.3	66	36	38
d(CG/GC)	–10.6	4.2	84	45	49
d(GC/CG)	–9.8	5.6	112	61	65
d(GG/CC)	–8.0	4.0	80	43	46
Initiation per GC	0.1	–10.2	–204	–110	–118
Initiation per AT	2.3	–5.2	–104	–56	–60

[a]Correction factor for self-complementary sequences is 0 kcal mol^{-1} for all cosolutes since it is independent of crowding environment.

Table S8. Nearest-neighbor parameters for $\Delta S^\circ_{[cation]}$ and $\Delta S^\circ_{[40\ wt\%\ PEG\ 200]}$ in 100 mM NaCl along with prefactors (m_{cs}) for different cosolutes[a].

Sequence	$\Delta S^\circ_{[cation]}$ (cal mol^{-1} K^{-1})	$\Delta S^\circ_{[40wt\%\ PEG\ 200]}$ (cal mol^{-1} K^{-1})	$m_{PEG/1,2\ DME}$ (cal mol^{-1} K^{-1})	$m_{EG/GOL}$ (cal mol^{-1} K^{-1})	$m_{1,3\ PDO/2\text{-}ME}$ (cal mol^{-1} K^{-1})
d(AA/TT)	–23.3	4.1	82	49	43
d(AT/TA)	–21.3	–8.1	–162	–97	–85
d(TA/AT)	–22.0	8.7	174	104	91
d(CA/GT)	–23.4	–15.4	–308	–185	–162
d(GT/CA)	–23.2	–3.6	–72	–43	–30
d(CT/GA)	–21.5	13.6	272	163	38
d(GA/CT)	–23.4	10.4	208	125	109
d(CG/GC)	–28.2	12.1	242	145	127
d(GC/CG)	–25.0	15.7	314	188	165
d(GG/CC)	–20.4	11.5	230	138	121
Initiation per GC	–2.8	–32.3	–646	–388	–339
Initiation per AT	4.1	–16.8	–336	–202	–176

[a]Correction factor for self-complementary sequences is -1.4 cal mol^{-1} K^{-1} for all cosolutes since it is independent of crowding environment.

Table S9. Nearest-neighbor parameters for $\Delta G°_{NN,\,[cation]}$ and $\Delta G°_{NN,\,[40\ wt\%\ PEG\ 200]}$ in 100 mM NaCl at 25 ºC along with prefactors (m_{cs})[a].

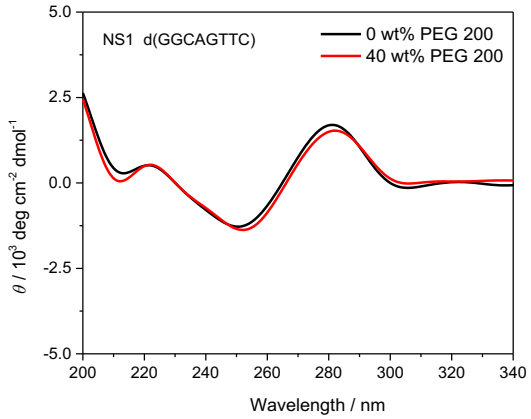
Sequence	$\Delta G°_{25\ NN,\,[cation]}$ (kcal mol^{-1})	$\Delta G°_{25\ NN,\,[40wt\%\ PEG\ 200]}$ (kcal mol^{-1})	$m_{PEG/1,2\ DME}$ (kcal mol^{-1})
d(AA/TT)	–0.95	0.18	3.6
d(AT/TA)	–0.85	0.22	4.4
d(TA/AT)	–0.64	0.31	6.2
d(CA/GT)	–1.52	–0.01	–0.2
d(GT/CA)	–1.48	0.27	5.5
d(CT/GA)	–1.39	0.35	6.9
d(GA/CT)	–1.22	0.20	4.0
d(CG/GC)	–2.19	0.59	11.8
d(GC/CG)	–2.35	0.92	18.4
d(GG/CC)	–1.92	0.57	11.4
Initiation per GC	0.93	–0.57	–11.4
Initiation per AT	1.08	–0.19	–3.8

[a]Correction factor for self-complementary sequences is 0.4 kcal mol^{-1} for all cosolutes since it is independent of the environment.

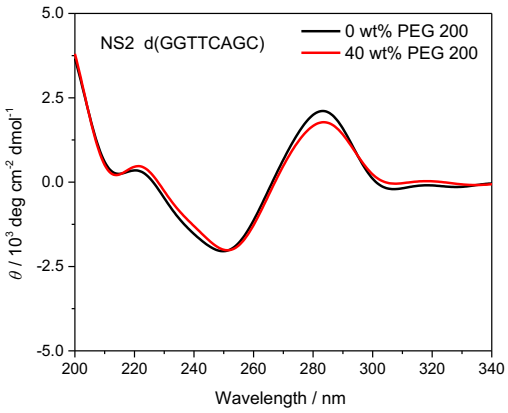
Table S10. Calculated $\Delta G°_{25,\ [crowder]}$ for $(ACTG)_3$ and $(ACTG)_4$ in various PEG 200 concentrations at 25 ºC in 100 mM NaCl.

Solution	$\Delta G°_{25,[crowder]}$ (kcal mol^{-1})	$\Delta G°_{25,[crowder]}$ (kcal mol^{-1})
	$d(ACTG)_3$	$d(ACTG)_4$
10 wt% PEG 200	0.35	0.54
20 wt% PEG 200	0.83	1.18
30 wt% PEG 200	1.21	1.71
40 wt% PEG 200	1.47	2.28
50 wt% PEG 200	2.01	2.83

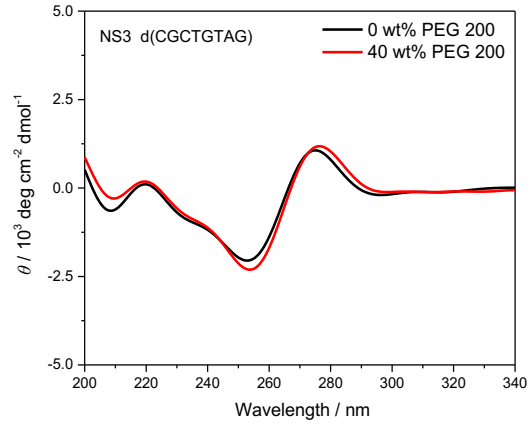
(1)



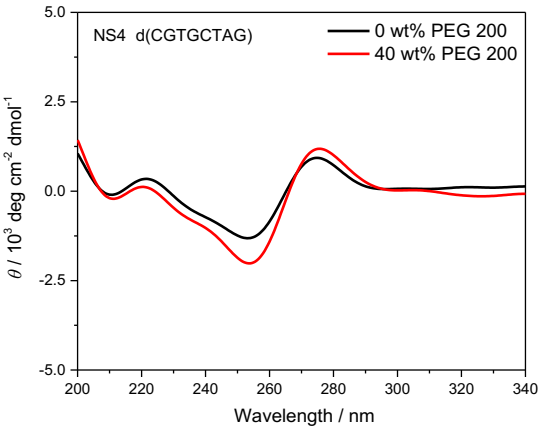
(2)



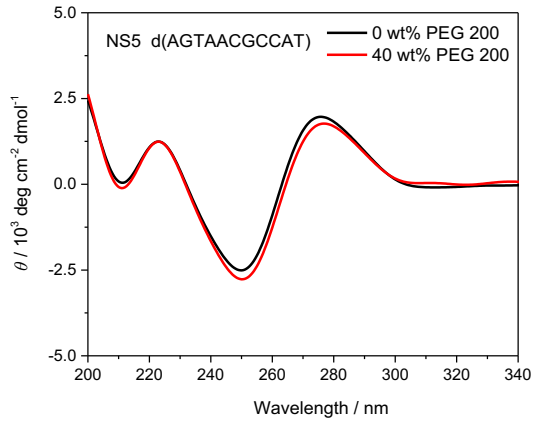
(3)



(4)



(5)



(6)

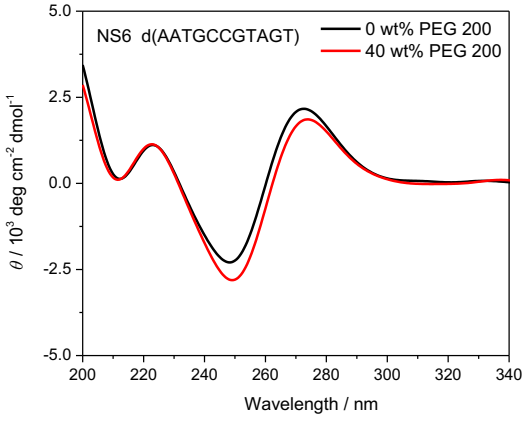
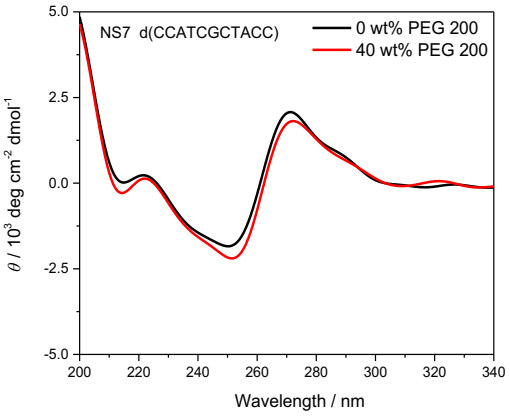
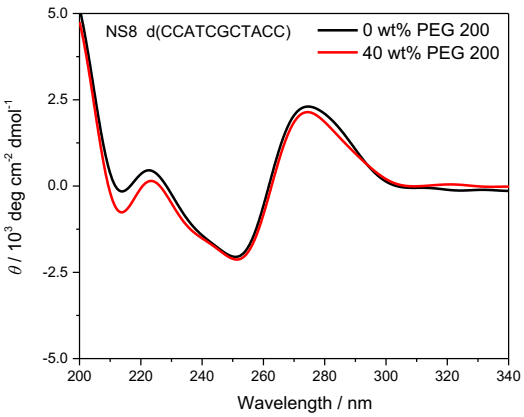


Figure S1 Continued

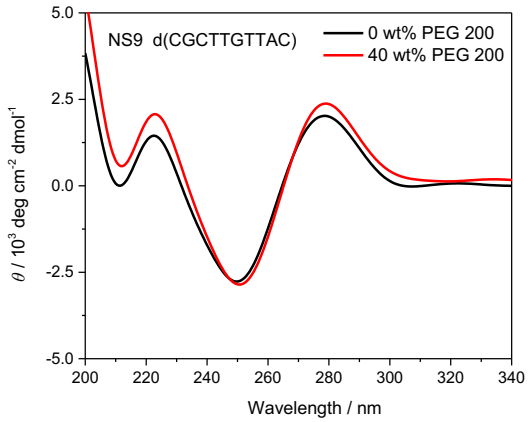
(7)



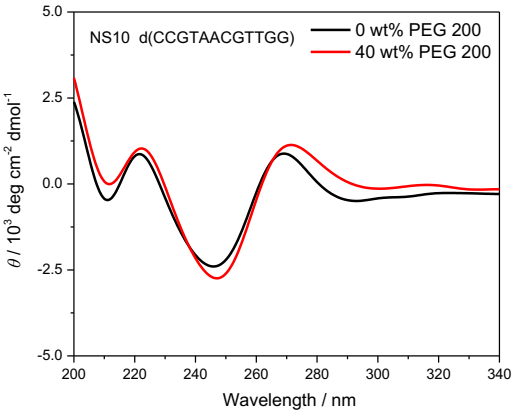
(8)



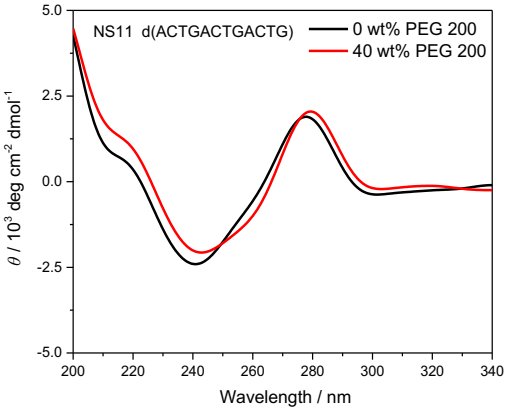
(9)



(10)



(11)



(12)

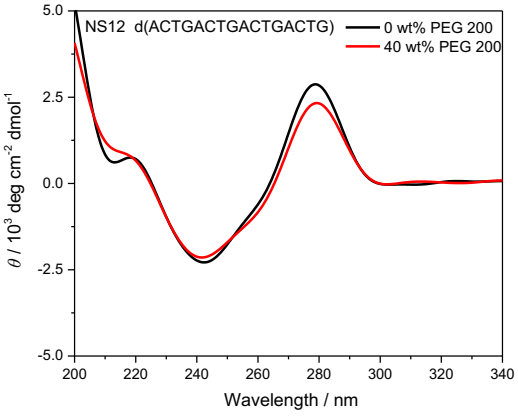
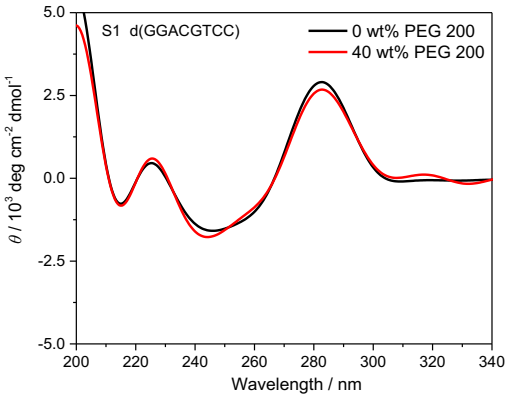
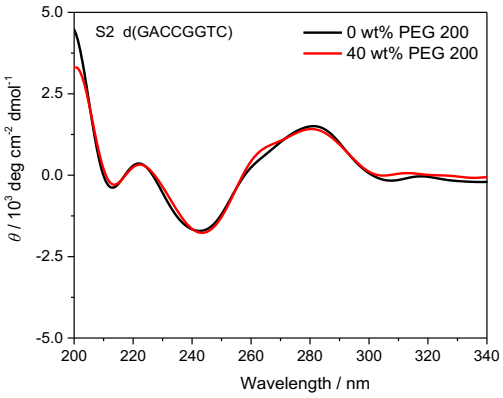


Figure S1 Continued

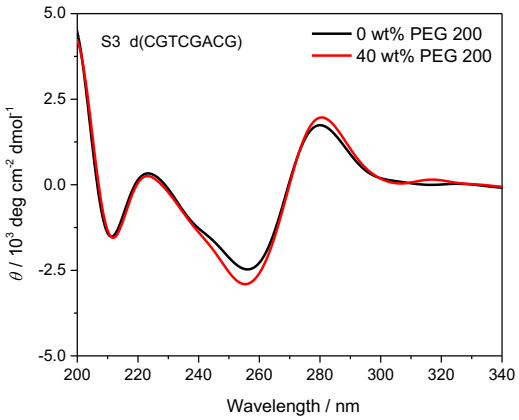
(13)



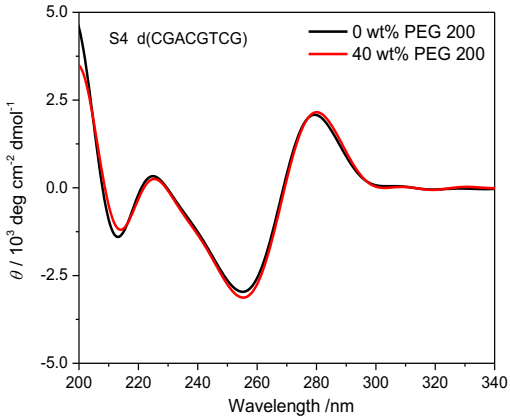
(14)



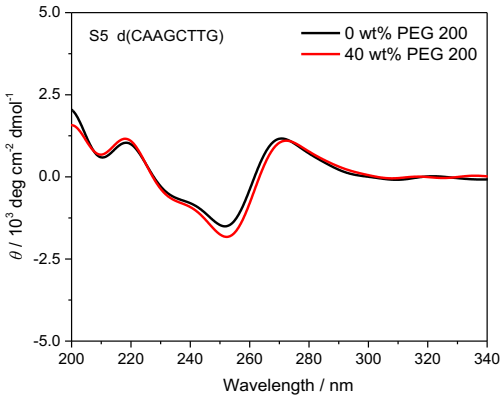
(15)



(16)



(17)



(18)

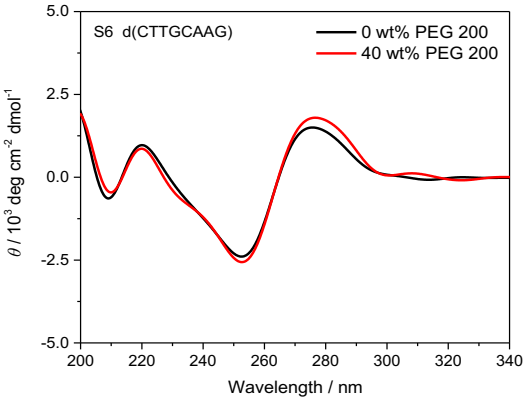
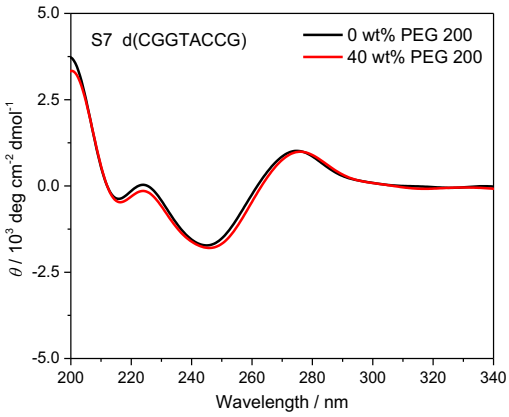
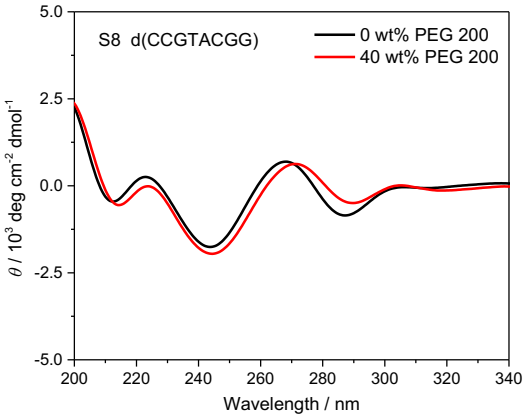


Figure S1 Continued

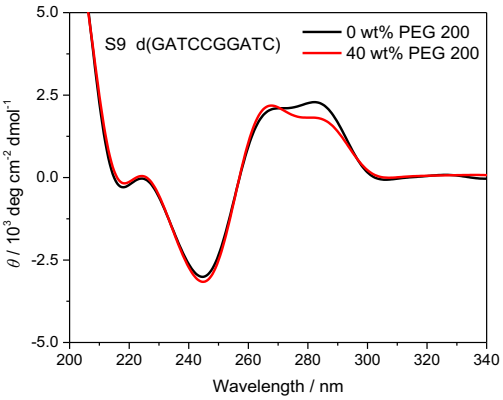
(19)



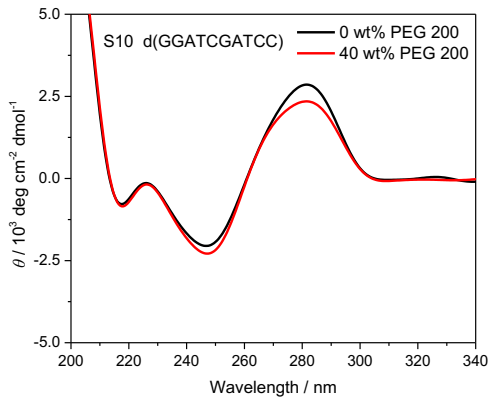
(20)



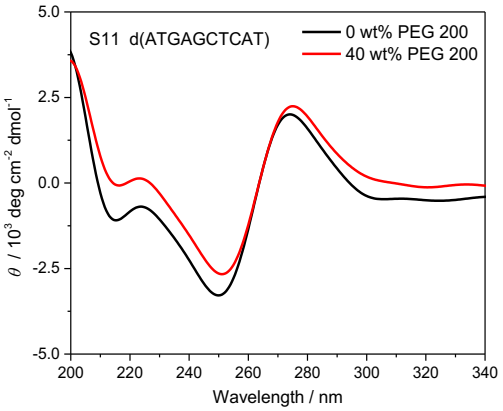
(21)



(22)



(23)



(24)

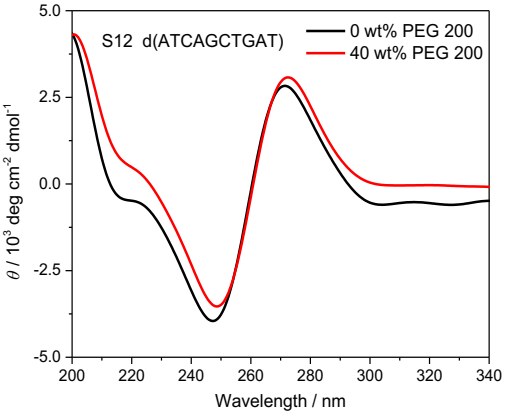
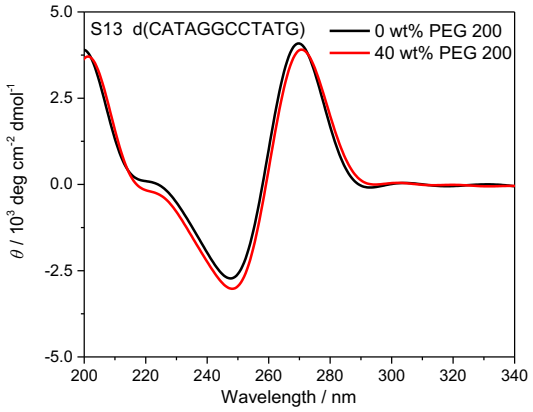
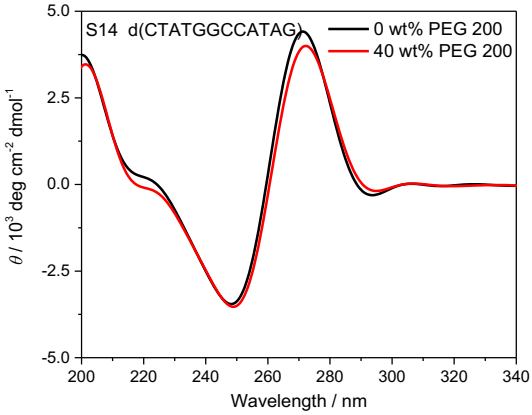


Figure S1 Continued

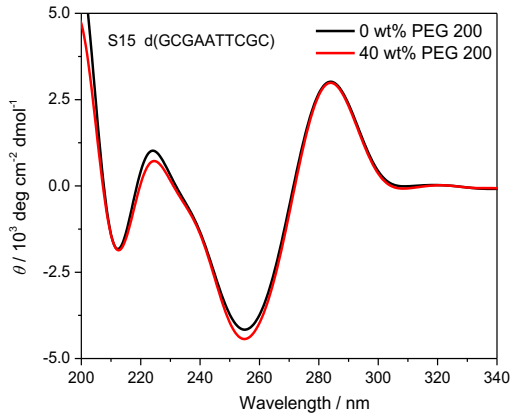
(25)



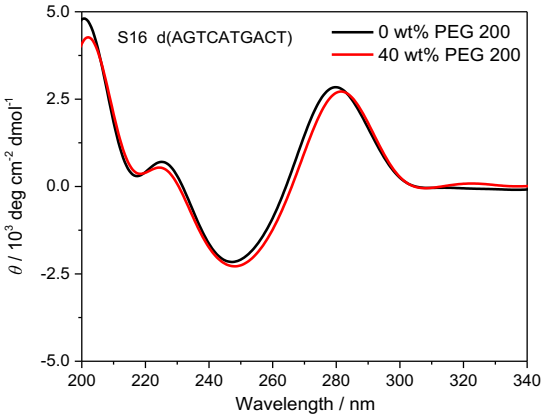
(26)



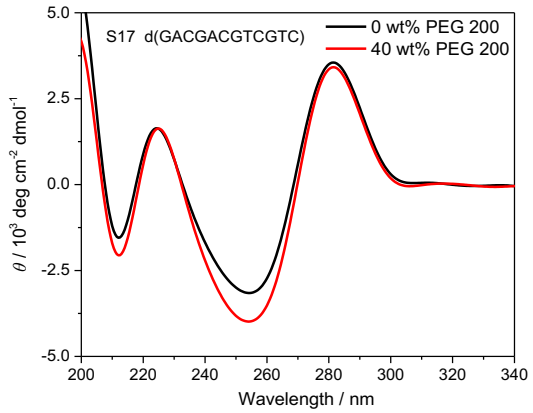
(27)



(28)



(29)



(30)

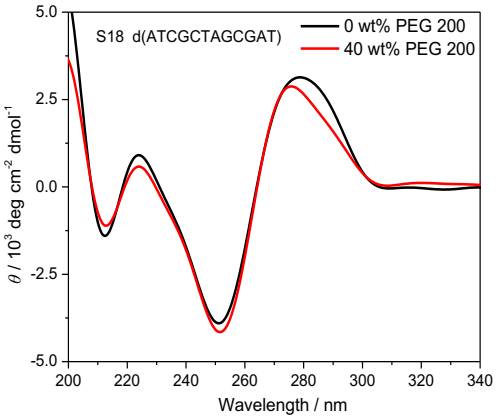
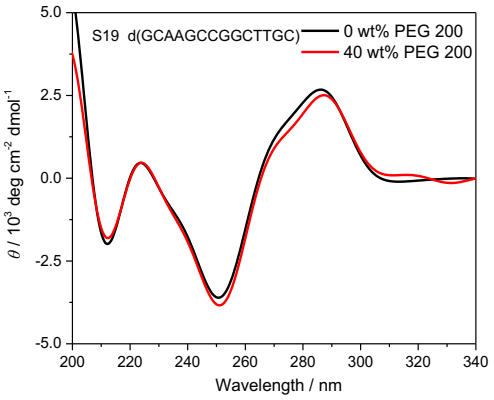
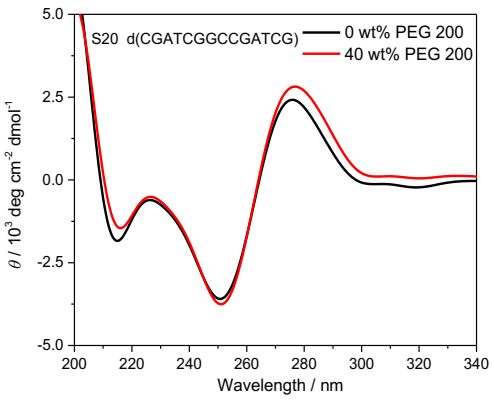


Figure S1 Continued

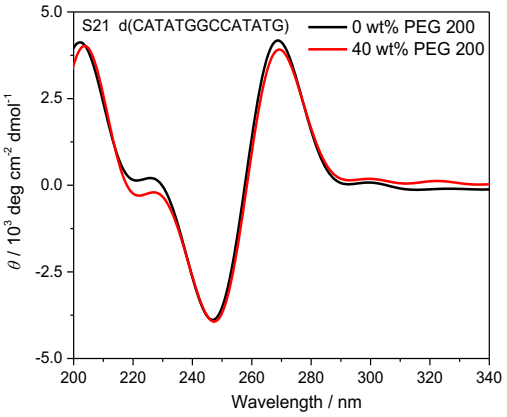
(31)



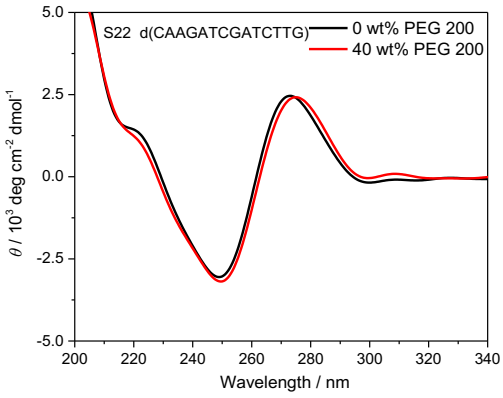
(32)



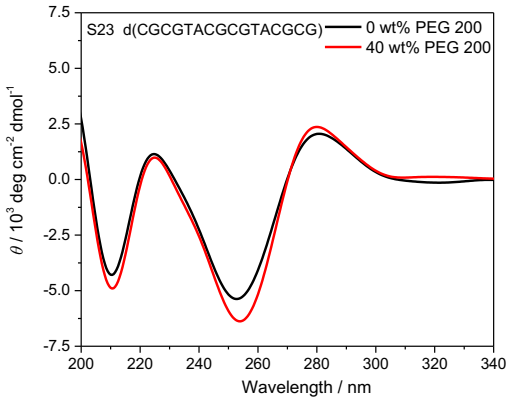
(33)



(34)



(35)



(36)

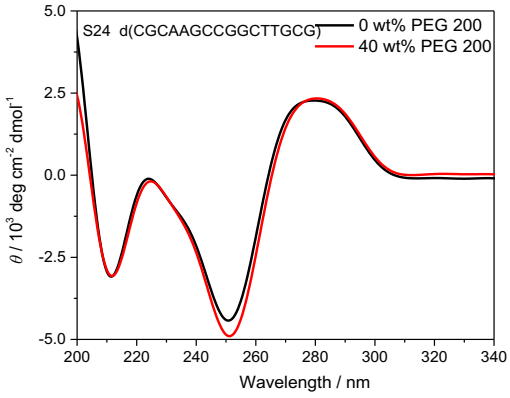
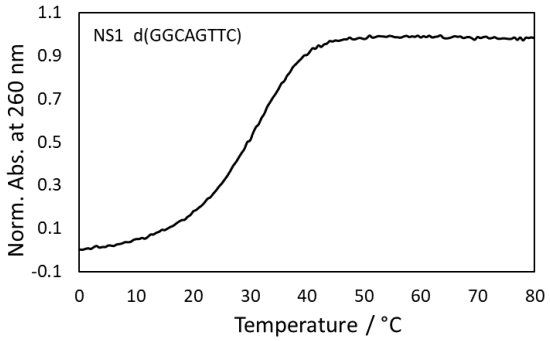


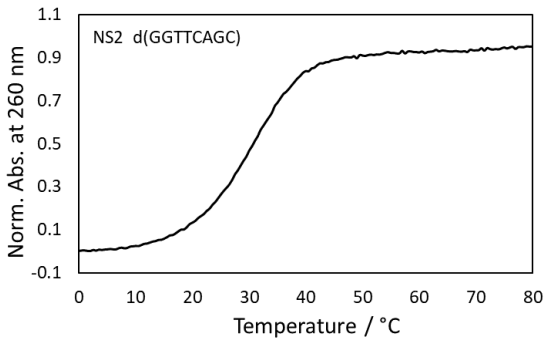
Figure S1 Continued

Figure S1. Circular dichroic spectra of 20 μM oligonucleotides at 4 °C in buffer containing 100 mM NaCl, 10 mM Na_2HPO_4 (pH 7.0) and 1 mM Na_2EDTA in absence (black) and presence (red) of 40 wt% PEG 200 for (1) d(GGCAGTTC) (No. NS1), (2) d(GGTTCAGC) (No. NS2), (3) d(CGCTGTAG) (No. NS3), (4) d(CGTGCTAG) (No. NS4), (5) d(AGTAACGCCAT) (No. NS5), (6) d(AATGCCGTAGT) (No.NS6), (7) d(CCATCGCTACC) (No. NS7), (8) d(CGATGGCCTAC) (No. NS8), (9) d(CGCTTGTTAC) (No. NS9), (10) (CCGTAACGTTGG) (No. NS10), (11) d(ACTGACTGACTG) (No. NS11), (12) d(ACTGACTGACTGACTG) (No. NS12), (13) d(GGACGTCC) (No. S1), (14) d(GACCGGTC) (No. S2), (15) d(CGTCGACG) (No. S3), (16) d(CGACGTCG) (No. S4), (17) d(CAAGCTTG) (No. S5), (18) d(CTTGCAAG) (No. S6), (19) d(CGGTACCG) (No. S7), (20) d(CCGTACGG) (No. S8), (21) d(GATCCGGATC) (No. S9), (22) d(GGATCGATCC) (No. S10), (23) d(ATGAGCTCAT) (No. S11),(24) d(ATCAGCTGAT) (No. S12), (25) d(CATAGGCCTATG) (No. S13), (26) d(CTATGGCCATAG) (No. S14), (27) d(GCGAATTCGC) (No. S15), (28) d(AGTCATGACT) (No. S16), (29) d(GACGACGTCGTC) (No. S17), (30) d(ATCGCTAGCGAT) (No. S18), (31) d(GCAAGCCGGCTTGC) (No. S19), (32) d(CGATCGGCCGATCG) (No. S20), (33) d(CATATGGCCATATG) (No. S21), (34) d(CAAGATCGATCTTG) (No. S22), (35) d(CGCGTACGCGTACGCG) (No. S23) and (36) d(CGCAAGCCGGCTTGCG) (No. S24).

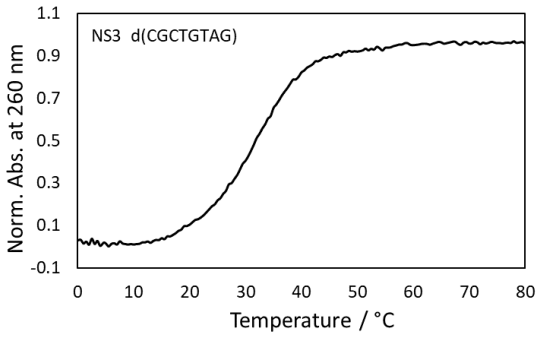
(1)



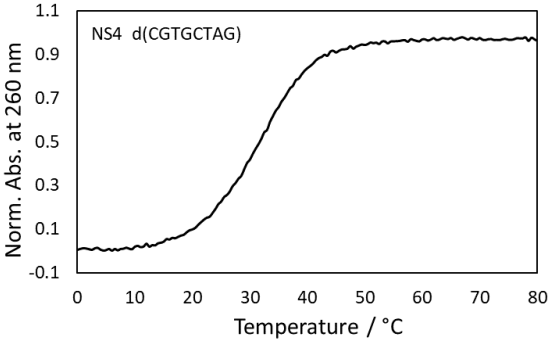
(2)



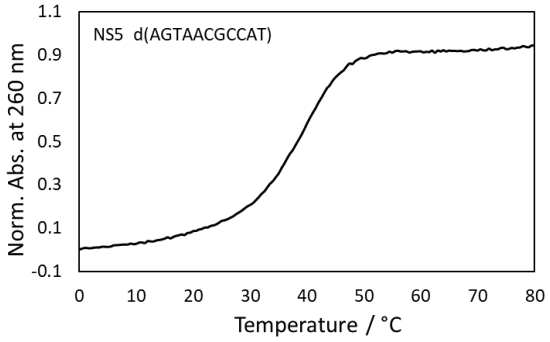
(3)



(4)



(5)



(6)

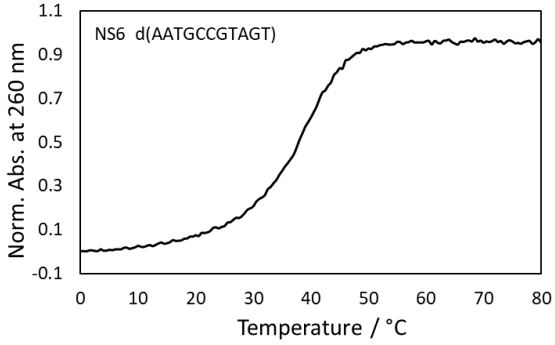
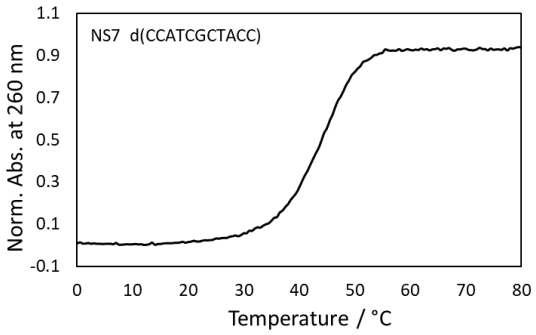
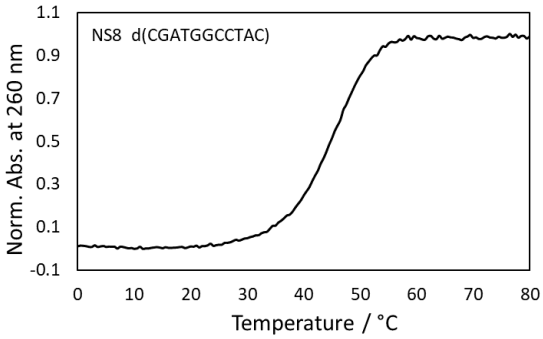


Figure S2 Continued

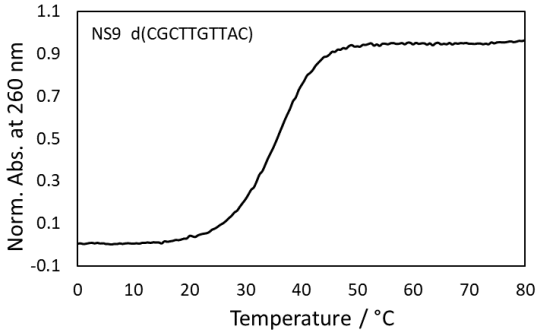
(7)



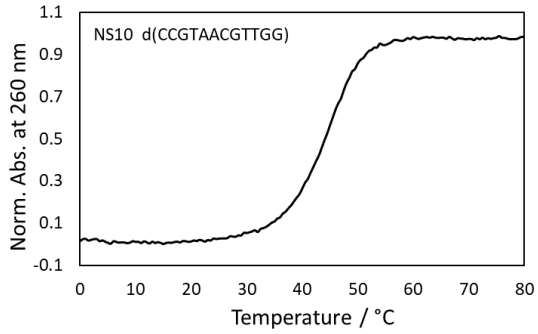
(8)



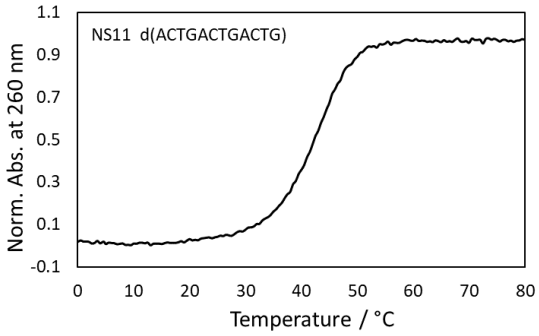
(9)



(10)



(11)



(12)

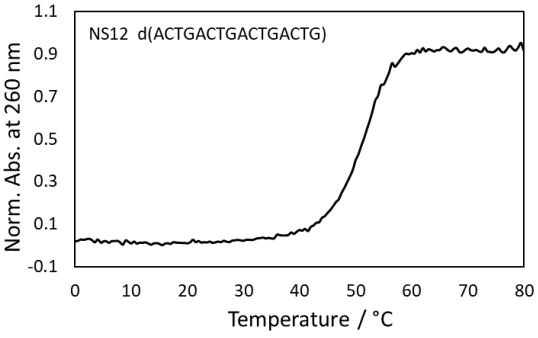
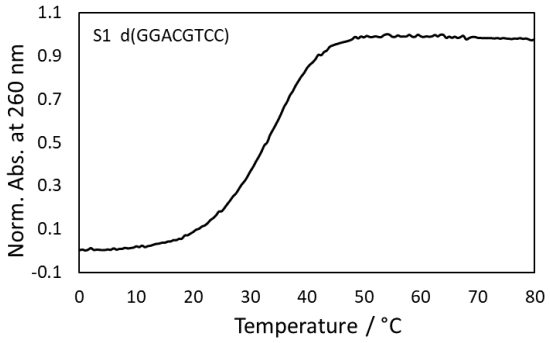
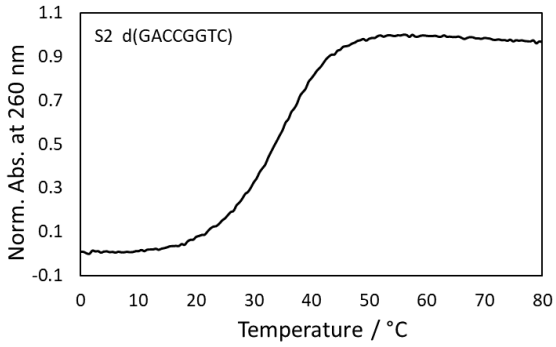


Figure S2 Continued

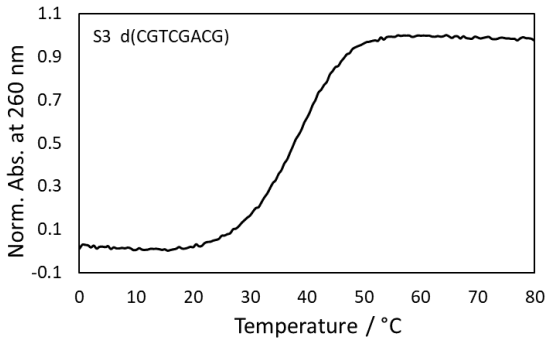
(13)



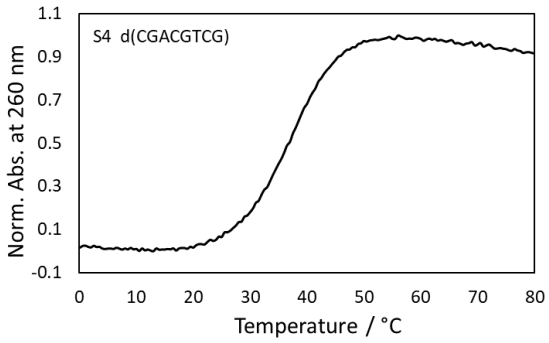
(14)



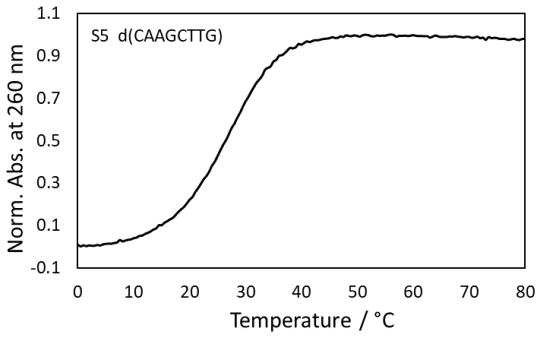
(15)



(16)



(17)



(18)

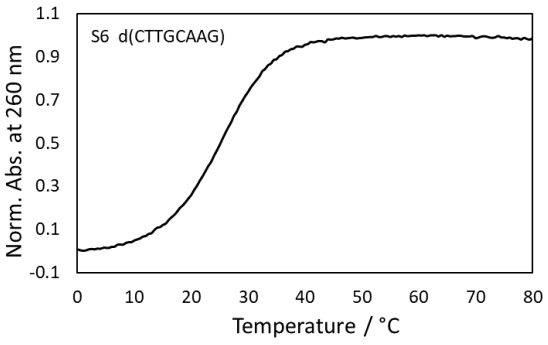
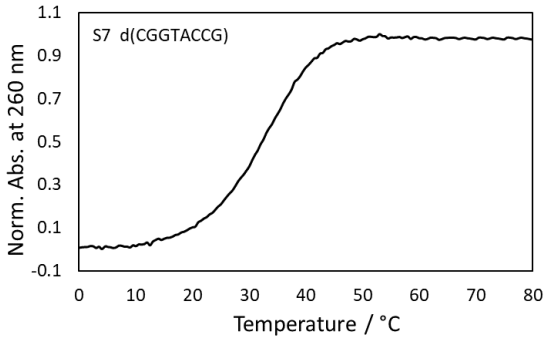
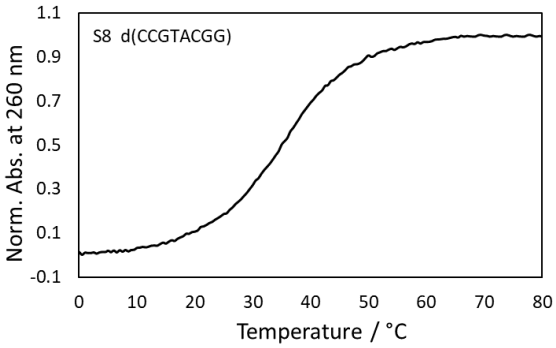


Figure S2 Continued

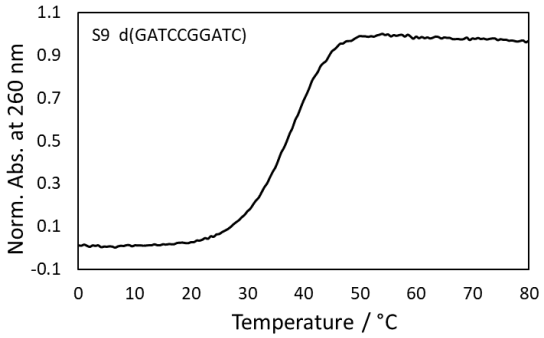
(19)



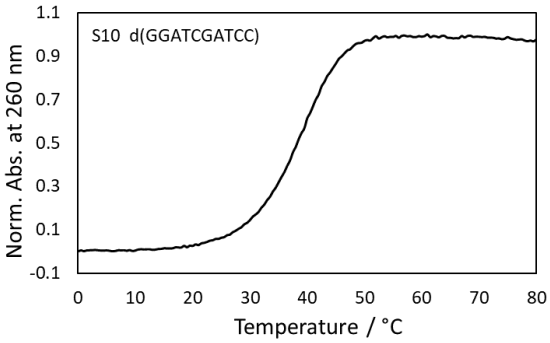
(20)



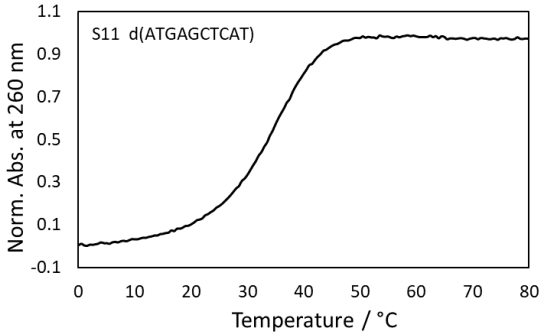
(21)



(22)



(23)



(24)

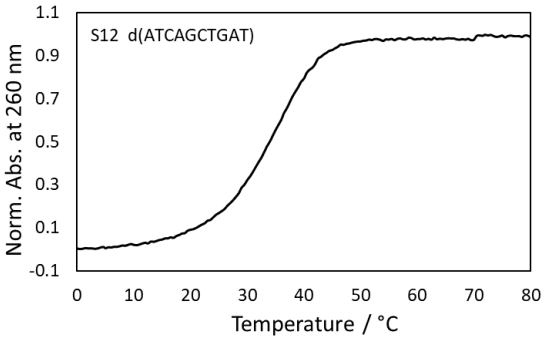
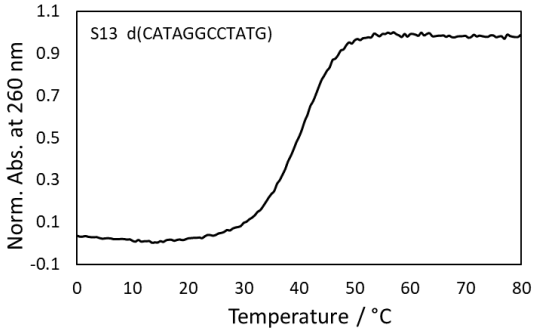
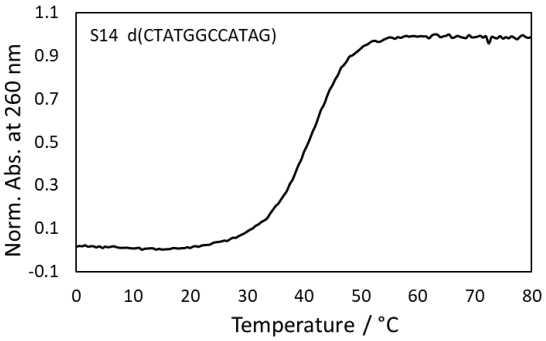


Figure S2 Continued

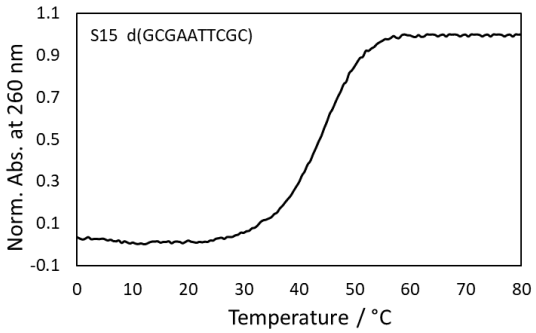
(25)



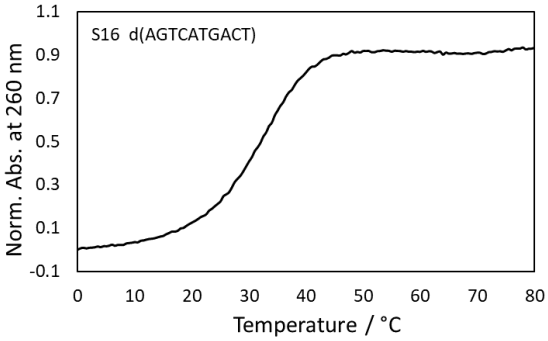
(26)



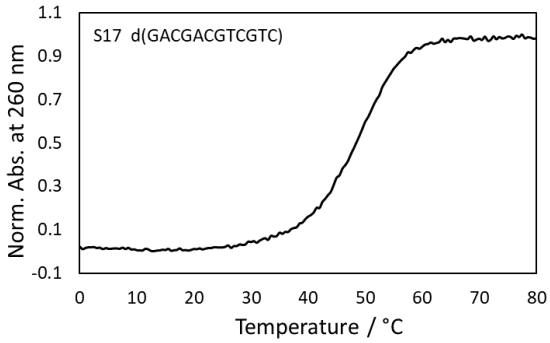
(27)



(28)



(29)



(30)

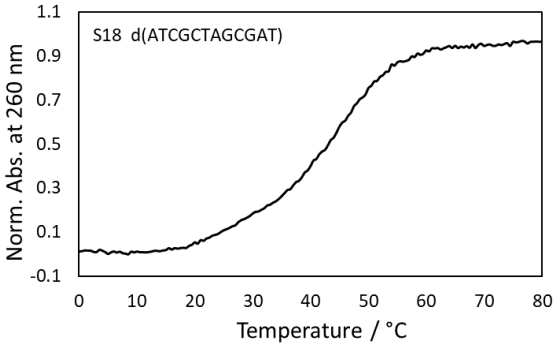


Figure S2 Continued

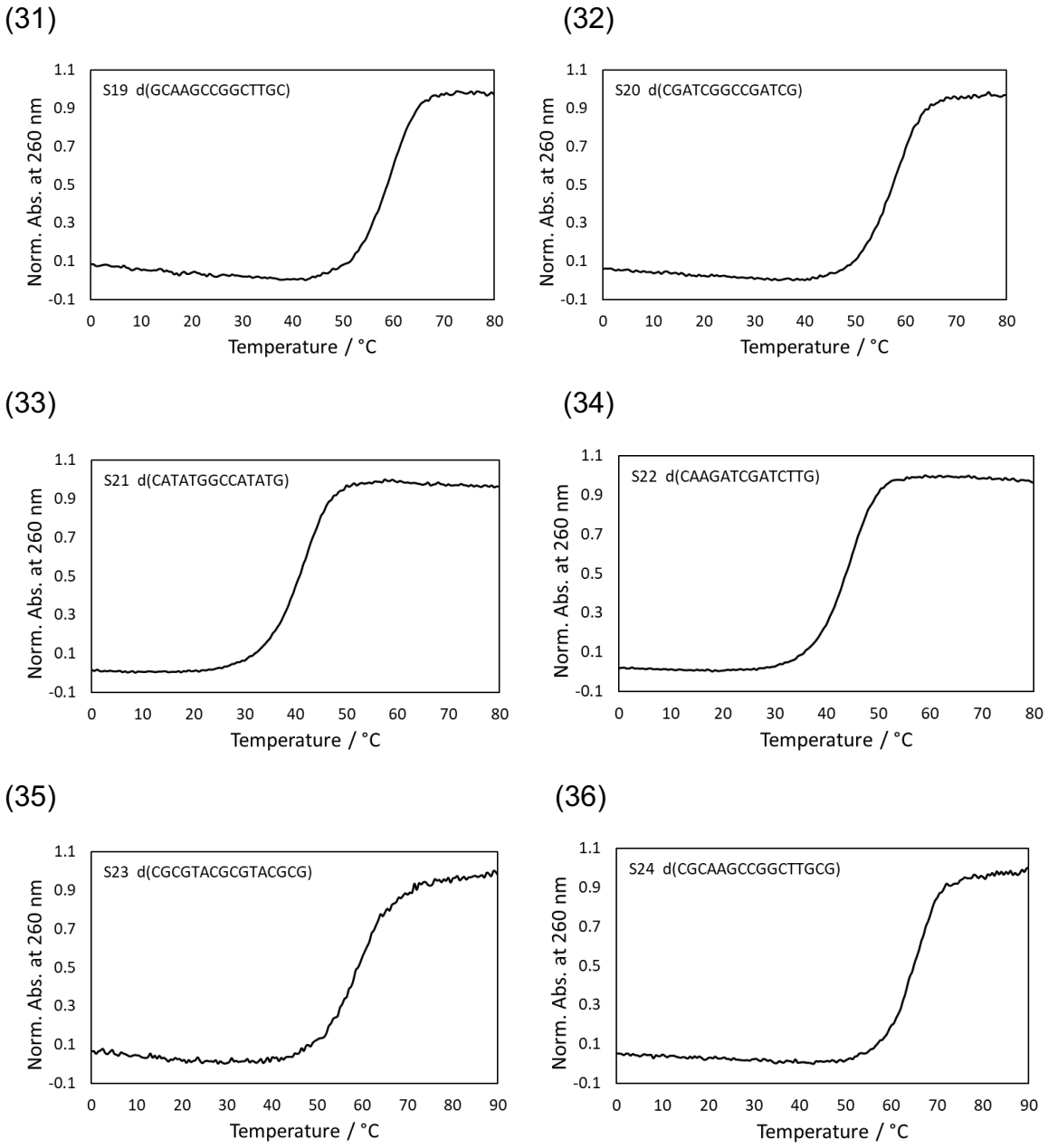
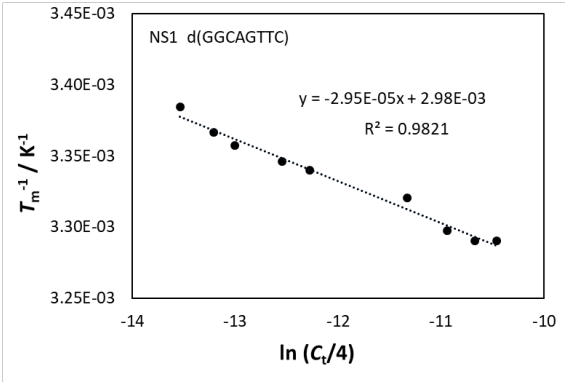


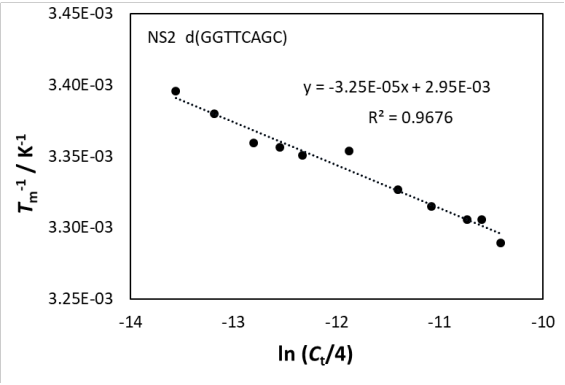
Figure S2. Representative UV melting curves of 100 μM oligonucleotides in 40 wt% PEG 200 with 100 mM NaCl for (1) d(GGCAGTTC) (No. NS1), (2) d(GGTTCAGC) (No. NS2), (3) d(CGCTGTAG) (No. NS3), (4) d(CGTGCTAG) (No. NS4), (5) d(AGTAACGCCAT) (No. NS5), (6) d(AATGCCGTAGT) (No.NS6), (7) d(CCATCGCTACC) (No. NS7), (8) d(CGATGGCCTAC) (No. NS8), (9) d(CGCTTGTTAC) (No. NS9), (10) (CCGTAACGTTGG) (No. NS10), (11) d(ACTGACTGACTG) (No. NS11), (12) d(ACTGACTGACTGACTG) (No. NS12), (13) d(GGACGTCC) (No. S1), (14) d(GACCGGTC) (No. S2), (15) d(CGTCGACG)

(No. S3), (16) d(CGACGTCG) (No. S4), (17) d(CAAGCTTG) (No. S5), (18) d(CTTGCAAG) (No. S6), (19) d(CGGTACCG) (No. S7), (20) d(CCGTACGG) (No. S8), (21) d(GATCCGGATC) (No. S9), (22) d(GGATCGATCC) (No. S10), (23) d(ATGAGCTCAT) (No. S11),(24) d(ATCAGCTGAT) (No. S12), (25) d(CATAGGCCTATG) (No. S13), (26) d(CTATGGCCATAG) (No. S14), (27) d(GCGAATTCGC) (No. S15), (28) d(AGTCATGACT) (No. S16), (29) d(GACGACGTCGTC) (No. S17), (30) d(ATCGCTAGCGAT) (No. S18), (31) d(GCAAGCCGGCTTGC) (No. S19), (32) d(CGATCGGCCGATCG) (No. S20), (33) d(CATATGGCCATATG) (No. S21), (34) d(CAAGATCGATCTTG) (No. S22), (35) d(CGCGTACGCGTACGCG) (No. S23) and (36) d(CGCAAGCCGGCTTGCG) (No. S24).

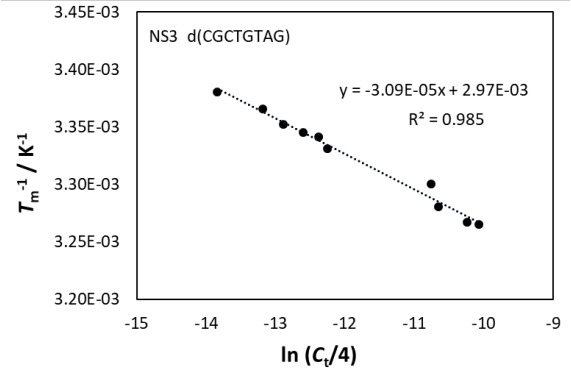
(1)



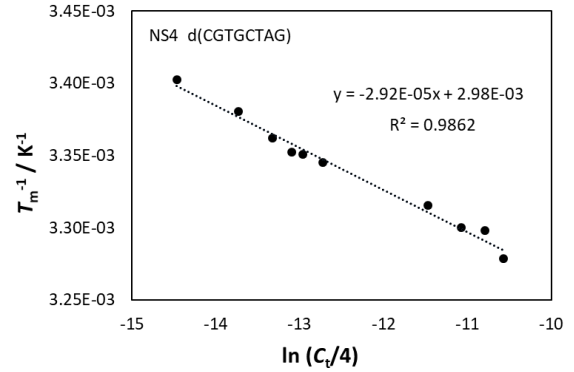
(2)



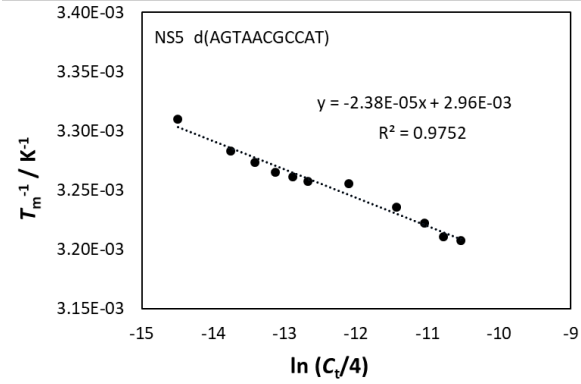
(3)



(4)



(5)



(6)

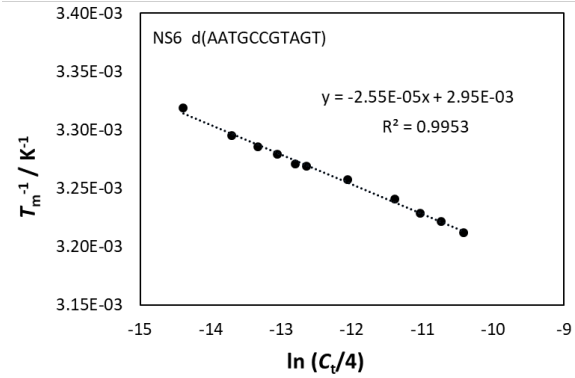
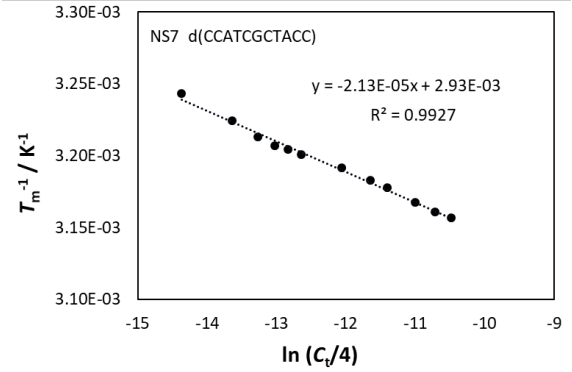
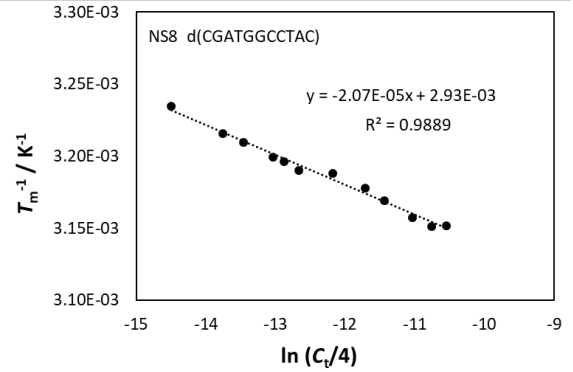


Figure S3 Continued

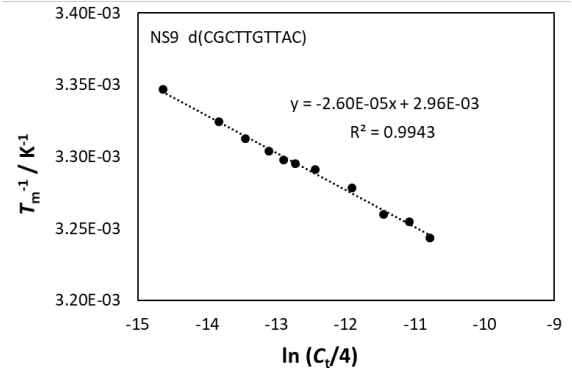
(7)



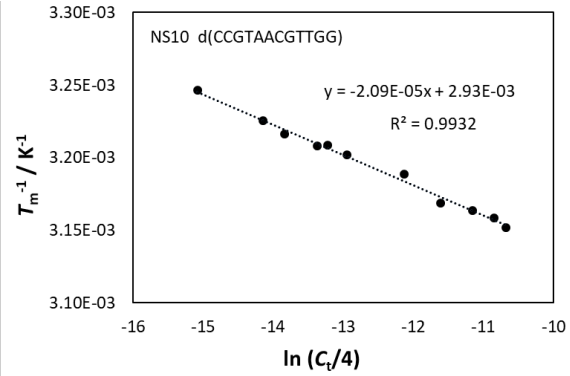
(8)



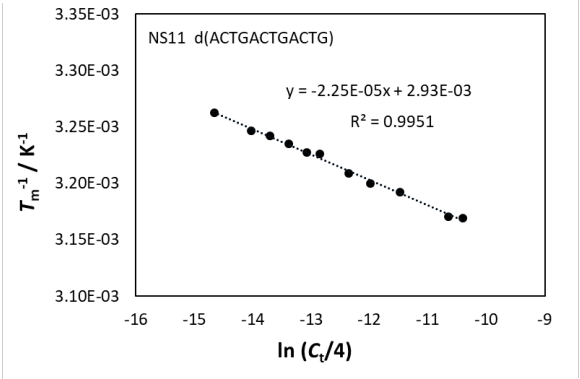
(9)



(10)



(11)



(12)

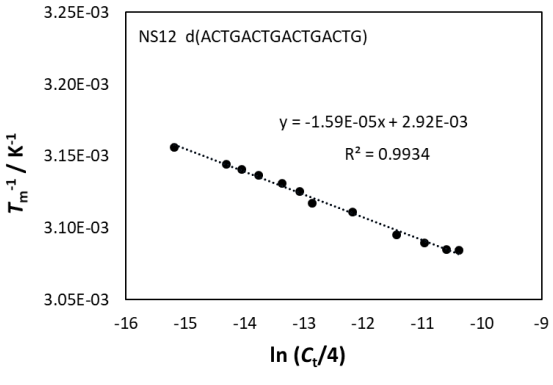
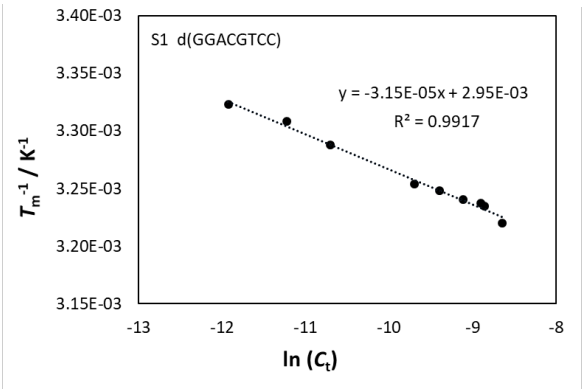
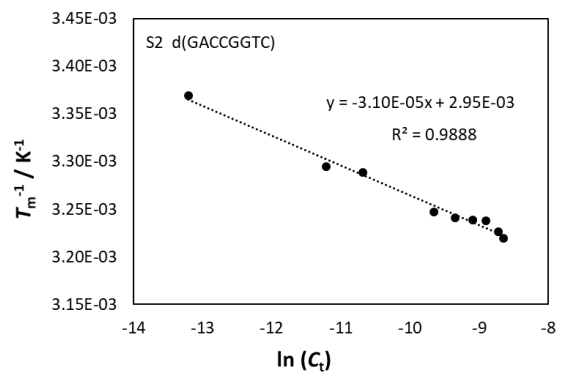


Figure S3 Continued

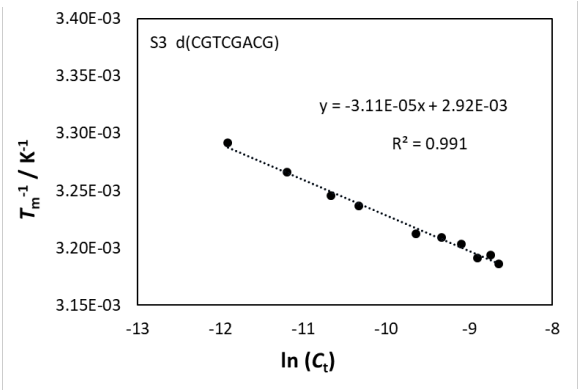
(13)



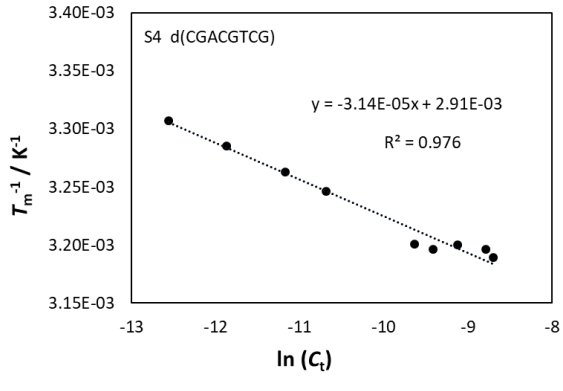
(14)



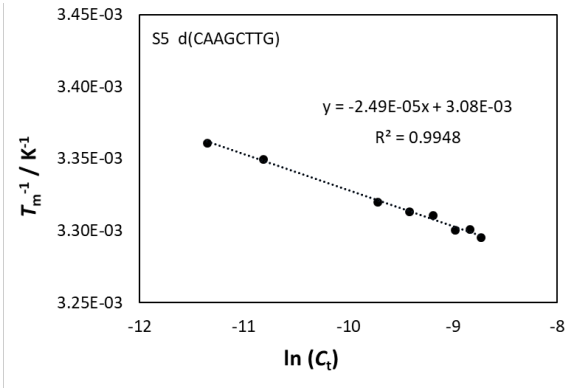
(15)



(16)



(17)



(18)

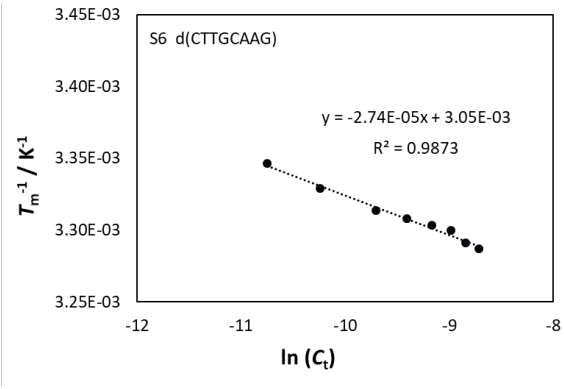
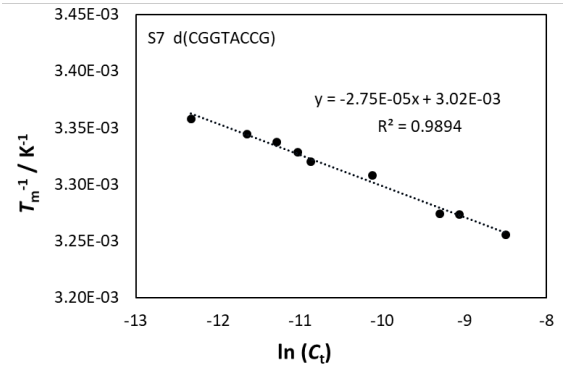
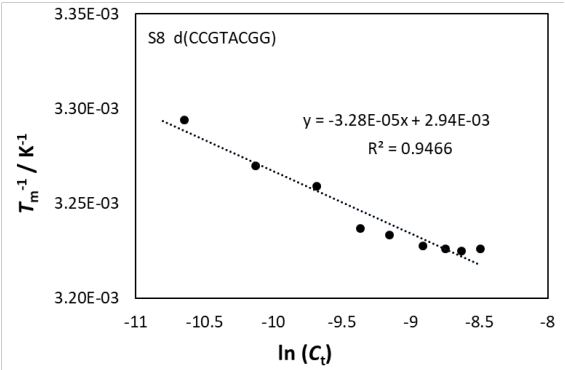


Figure S3 Continued

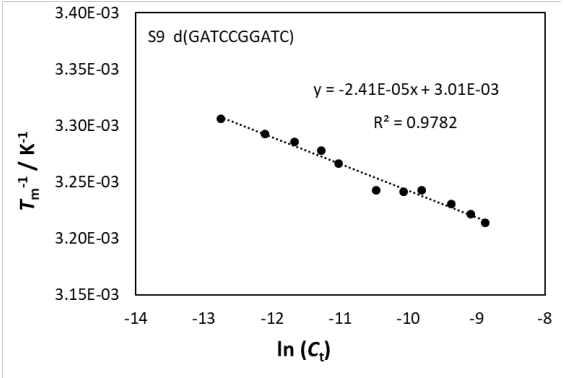
(19)



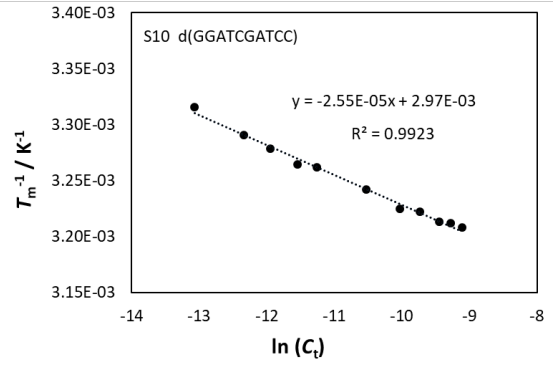
(20)



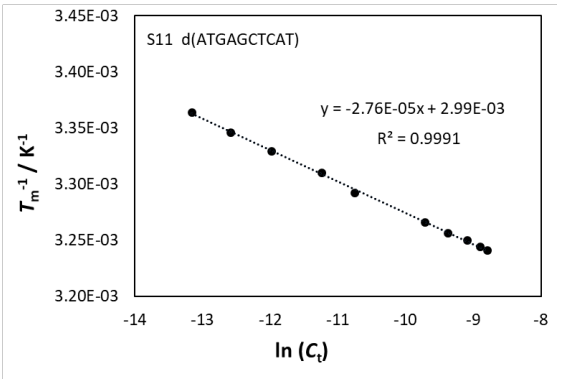
(21)



(22)



(23)



(24)

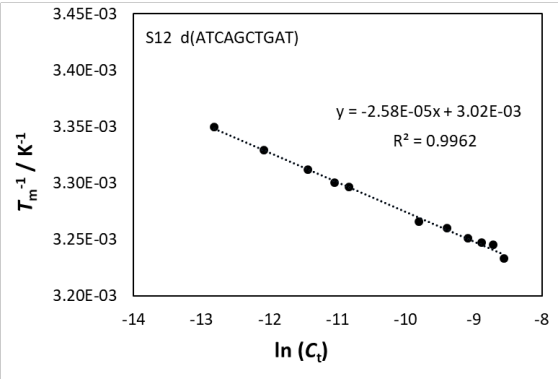
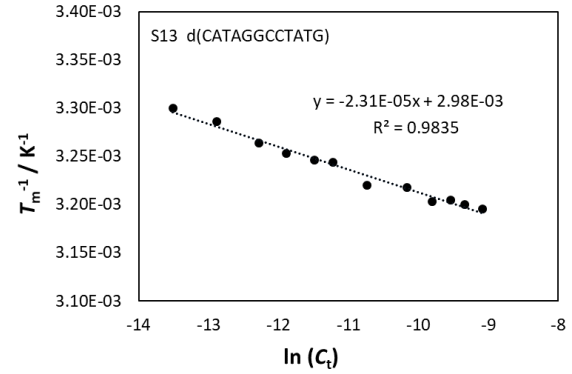
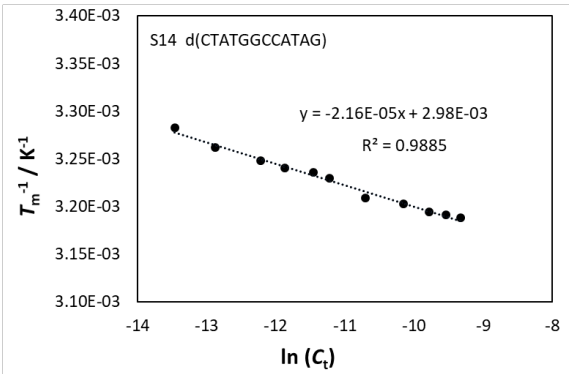


Figure S3 Continued

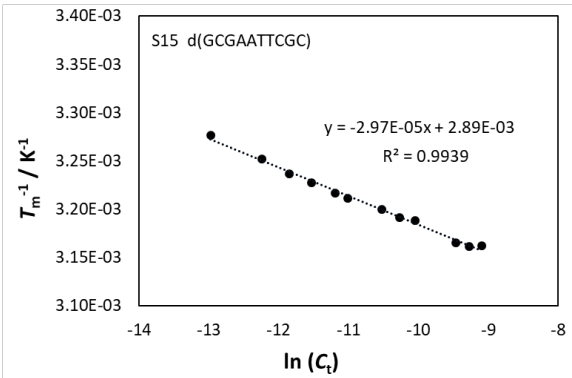
(25)



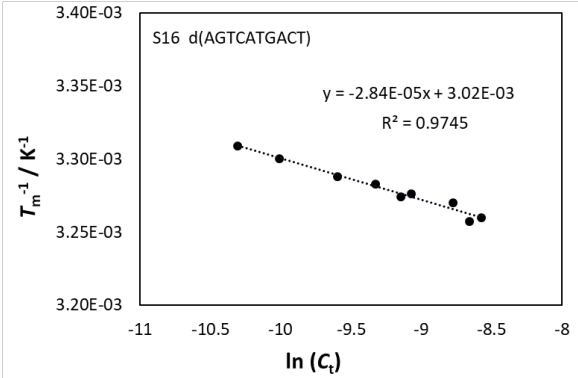
(26)



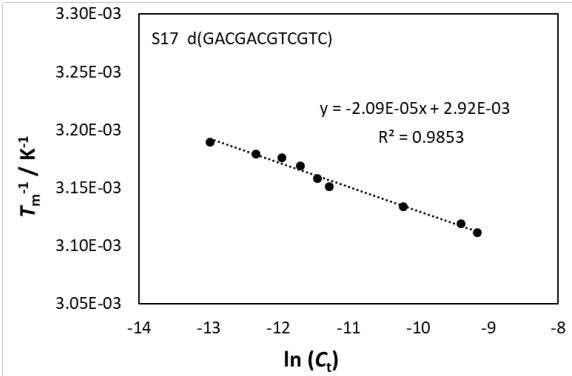
(27)



(28)



(29)



(30)

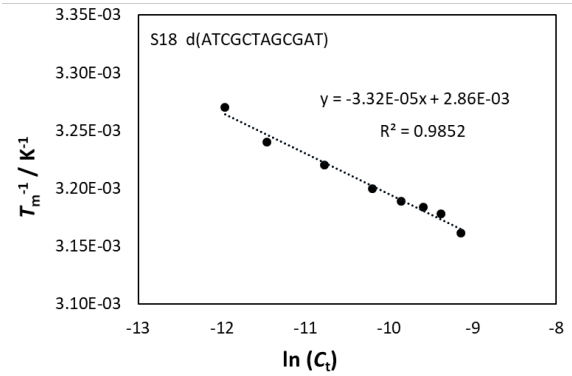


Figure S3 Continued

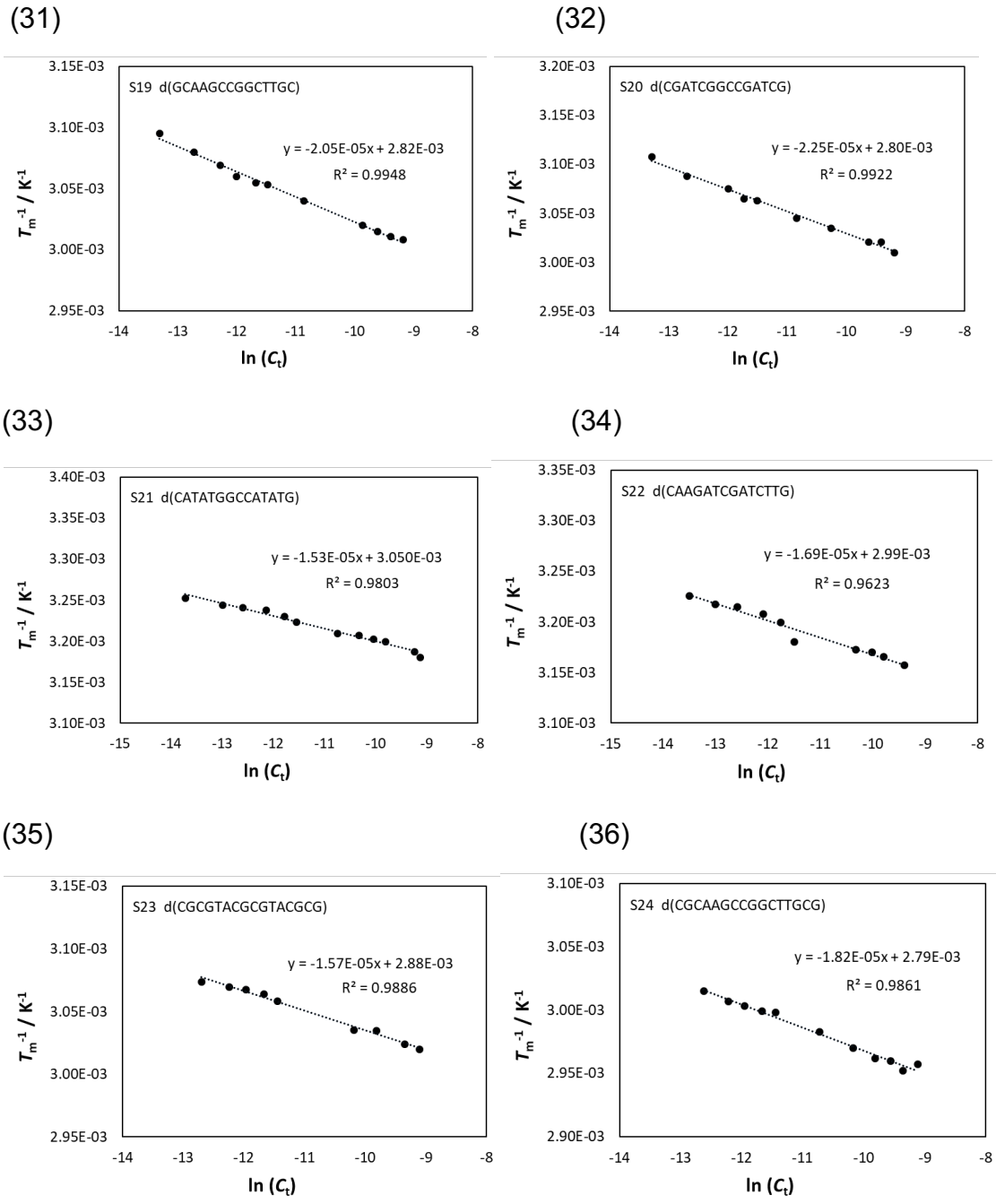


Figure S3. T_m^{-1} vs. ln (C_t/s) plots in 40 wt% PEG 200 with 100 mM NaCl for (1) d(GGCAGTTC) (No. NS1), (2) d(GGTTCAGC) (No. NS2), (3) d(CGCTGTAG) (No. NS3), (4) d(CGTGCTAG) (No. NS4), (5) d(AGTAACGCCAT) (No. NS5), (6) d(AATGCCGTAGT) (No.NS6), (7) d(CCATCGCTACC) (No. NS7), (8) d(CGATGGCCTAC) (No. NS8), (9) d(CGCTTGTTAC) (No. NS9), (10) (CCGTAACGTTGG) (No. NS10), (11)

d(ACTGACTGACTG) (No. NS11), (12) d(ACTGACTGACTGACTG) (No. NS12), (13) d(GGACGTCC) (No. S1), (14) d(GACCGGTC) (No. S2), (15) d(CGTCGACG) (No. S3), (16) d(CGACGTCG) (No. S4), (17) d(CAAGCTTG) (No. S5), (18) d(CTTGCAAG) (No. S6), (19) d(CGGTACCG) (No. S7), (20) d(CCGTACGG) (No. S8), (21) d(GATCCGGATC) (No. S9), (22) d(GGATCGATCC) (No. S10), (23) d(ATGAGCTCAT) (No. S11),(24) d(ATCAGCTGAT) (No. S12), (25) d(CATAGGCCTATG) (No. S13), (26) d(CTATGGCCATAG) (No. S14), (27) d(GCGAATTCGC) (No. S15), (28) d(AGTCATGACT) (No. S16), (29) d(GACGACGTCGTC) (No. S17), (30) d(ATCGCTAGCGAT) (No. S18), (31) d(GCAAGCCGGCTTGC) (No. S19), (32) d(CGATCGGCCGATCG) (No. S20), (33) d(CATATGGCCATATG) (No. S21), (34) d(CAAGATCGATCTTG) (No. S22), (35) d(CGCGTACGCGTACGCG) (No. S23) and (36) d(CGCAAGCCGGCTTGCG) (No. S24).

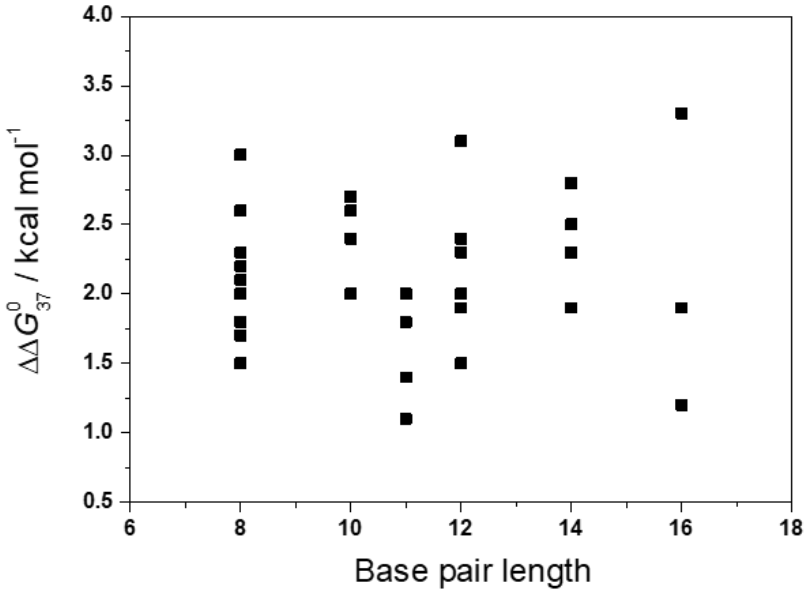


Figure S4. Destabilization ($\Delta\Delta G°_{37}$) of DNA duplexes in 40 wt% PEG 200 vs. base pair length of oligonucleotide duplexes investigated herein. Data were collected from Table 1.

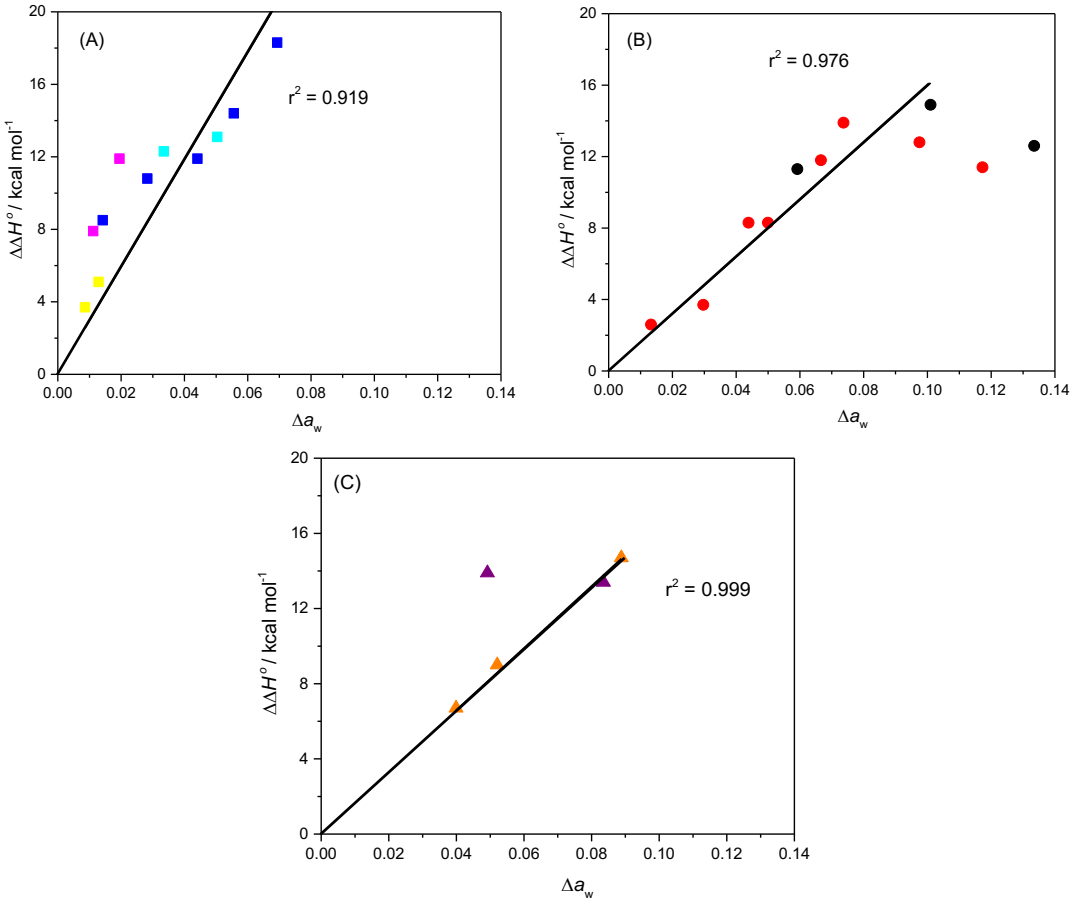


Figure S5. Differences in $\Delta H°$ between crowded conditions and no cosolute ($\Delta\Delta H°$) against Δa_w for d(ATGCGCAT), in 1 M NaCl, in various concentrations of 1, 2-dimethoxy ethane (cyan squares), PEG 200 (blue squares), PEG 2000 (magenta squares) and PEG 8000 (yellow squares) (A), ethylene glycol (red circles) and glycerol (black circles)(B) and 1,3-propanediol (purple triangles) and 2-methoxyethanol (orange triangles) (C).

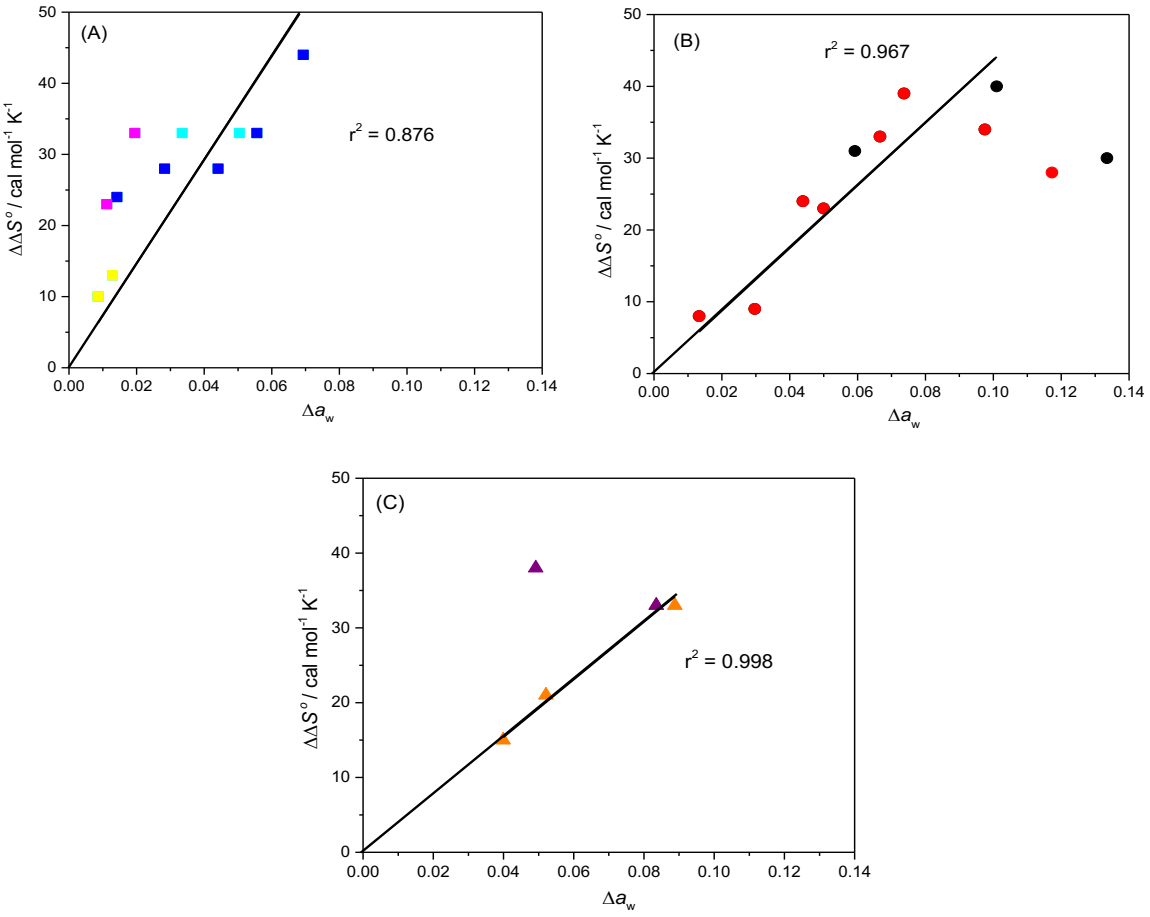


Figure S6. Differences in $\Delta S°$ between crowded conditions and without cosolute ($\Delta\Delta S°$) vs. Δa_w for d(ATGCGCAT), in 1 M NaCl, with various concentrations of 1, 2-dimethoxy ethane (cyan squares), PEG 200 (blue squares), PEG 2000 (magenta squares) and PEG 8000 (yellow squares) (A), ethylene glycol (red circles) and glycerol (black circles) (B) and 1,3-propanediol (purple triangles) and 2-methoxyethanol (orange triangles)(C).

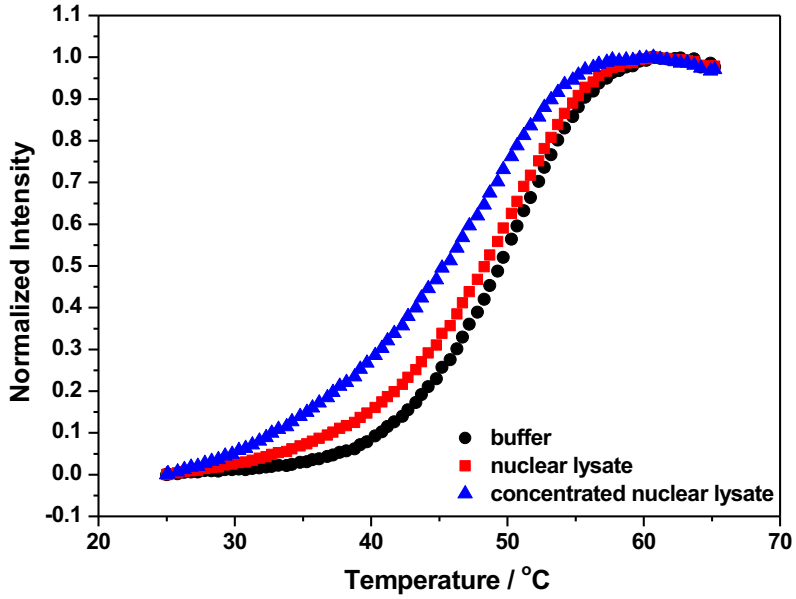


Figure S7. Melting assays using fluorescence-labeled d(AGTAACGCCAT) in HeLa cell nuclear lysates. 5´ end of AGTAACGCCAT was labelled with 6FAM and the 3´ end of the complementary sequence with quencher BHQ1. Annealed duplex (5 μM) was added to the solutions and fluorescence intensity of 6FAM was monitored with increasing temperature.

References

1. Sugimoto N, Satoh N, Yasuda K, & Nakano S (2001) Stabilization factors affecting duplex formation of peptide nucleic acid with DNA. *Biochemistry* 40(29):8444-8451.
2. Ghosh S*, et al.* (2019) Validation of the nearest-neighbor model for Watson-Crick self-complementary DNA duplexes in molecular crowding condition. *Nucleic Acids Res* 47(7):3284-3294.
3. SantaLucia J, Jr., Allawi HT, & Seneviratne PA (1996) Improved nearest-neighbor parameters for predicting DNA duplex stability. *Biochemistry* 35(11):3555-3562.
4. Sugimoto N, Nakano S, Yoneyama M, & Honda K (1996) Improved thermodynamic parameters and helix initiation factor to predict stability of DNA duplexes. *Nucleic Acids Res* 24(22):4501-4505.
5. Freier SM*, et al.* (1986) Improved free-energy parameters for predictions of RNA duplex stability. *Proc Natl Acad Sci USA* 83(24):9373-9377.
6. Xia T*, et al.* (1998) Thermodynamic parameters for an expanded nearest-neighbor model for formation of RNA duplexes with Watson-Crick base pairs. *Biochemistry* 37(42):14719-14735.
7. Sugimoto N*, et al.* (1995) Thermodynamic parameters to predict stability of RNA/DNA hybrid duplexes. *Biochemistry* 34(35):11211-11216.
8. Erie D, Sinha N, Olson W, Jones R, & Breslauer K (1987) A dumbbell-shaped, double-hairpin structure of DNA: a thermodynamic investigation. *Biochemistry* 26(22):7150-7159.
9. Rentzeperis D, Ho J, & Marky LA (1993) Contribution of loops and nicks to the formation of DNA dumbbells: melting behavior and ligand binding. *Biochemistry* 32(10):2564-2572.
10. SantaLucia J, Jr. (1998) A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proc Natl Acad Sci USA* 95(4):1460-1465.
11. Goobes R, Kahana N, Cohen O, & Minsky A (2003) Metabolic buffering exerted by macromolecular crowding on DNA-DNA interactions: origin and physiological significance. *Biochemistry* 42(8):2431-2440.
12. Fujiwara K & Nomura S-iM (2013) Condensation of an Additive-Free Cell Extract to Mimic the Conditions of Live Cells. *PLOS ONE* 8(1):e54155.
13. Huguet JM, Ribezzi-Crivellari M, Bizarro CV, & Ritort F (2017) Derivation of nearest-neighbor DNA parameters in magnesium from single molecule experiments. *Nucleic Acids Res* 45(22):12921-12931.
14. Nakano S, Karimata H, Ohmichi T, Kawakami J, & Sugimoto N (2004) The effect of molecular crowding with nucleotide length and cosolute structure on DNA duplex stability. *J Am Chem Soc* 126(44):14330-14331.