# Thermally activated irreversible homogenization of G-quadruplexes in an ALS/FTD-associated nucleotide expansion

Daniel Ross,[1,*] Olivia Lewis,[1] Olivia McLean,[1] Sundeep Bhanot,[1] Shane Donahue,[1] Rachael Baker,[1] Randi Dias,[1] David Eagerton,[1] Vaibhav Mohanty,[2,3,4,*] and Bidyut K. Mohanty[1,*]

[1] *Edward Via College of Osteopathic Medicine–Carolinas Campus, Spartanburg, SC 29303*
[2] *Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA 02138*
[3] *Harvard/MIT MD-PhD Program, Harvard Medical School, Boston, MA 02115 and Massachusetts Institute of Technology, Cambridge, MA 02139*
[4] *Program for Health Sciences and Technology, Harvard Medical School, Boston, MA 02115 and Massachusetts Institute of Technology, Cambridge, MA 02139*

A significant proportion of familial amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD) cases exhibit a substantial copy number expansion of the hexanucleotide GGGGCC/GGCCCC sequence in the *C9ORF72* gene. The GGGGCC sequence forms a non-canonical DNA structure called a G-quadruplex (G4) which has been associated with the disease states and with nucleic acid condensate formation, but there remains poor understanding of why the hexanucleotide repeat expansion occurs. An important clue may lie in the thermodynamic properties of various G4 configurations. Here, we use temperature-swept CD spectroscopy to observe configurational homogenization of an initially heterogeneous population of G4s over a small range of temperatures, demonstrating phase transition-like behavior. We further show that this reaction is irreversible, suggesting that the metastable non-parallel G4s are trapped over long timescales by a transition state barrier; upon heating, the relative height of the free energy barrier is lowered, and a thermally activated metastable-to-stable transition to parallel G4s rapidly occurs. We provide an analytical theory based on a two-state thermodynamic model which offers mechanistic insight into the experimental evidence. These findings suggest that kinetic regulation of non-canonical DNA structures may play a role in cellular homeostasis or disease pathogenesis.

## I. INTRODUCTION

Amyotrophic lateral sclerosis (ALS, Lou Gehrig's disease) and frontotemporal dementia (FTD) are devastating and generally fatal neuromotor diseases that, respectively, cause muscle weakness and brain atrophy, substantially increasing morbidity and mortality of affected individuals [1–5]. Some of the shared symptoms of ALS and FTD are changes in behavior, personality, and language, and persons with a spectrum of behavioral disturbances common to ALS and FTD form a clinical subtype known as ALS-FTD [5, 6]. Presently, there are few treatments for these disorders, which at best may delay some of the disease symptoms [3, 6, 7].

ALS, FTD, and ALS-FTD can be sporadic or familial. Various genetic mutations have been identified in patients with familial ALS, FTD, or ALS-FTD, many of which involve mutations in the *C9ORF72* gene [2, 6–9]. Intron 1 of *C9ORF72* typically contains 2-22 copies of the hexanucleotide repeat GGGGCC/GGCCCC which can expand to 700-1600 copies in pathological settings [5, 8, 9]. Transcription and downstream translation of these hexanucleotide repeat expansions (HREs) are key contributors to the pathogenesis of *C9ORF72*-associated ALS/FTD [10–12]. Although current research on development of ALS/FTD therapeutics often focuses on RNA and translated protein toxicity, the origin of the HRE and ALS/FTD lies solely in DNA and therefore targeting the mechanism of HRE is of equal importance [3].

Although the *C9ORF72* gene was discovered in 2011 [9], the core mechanism of its HRE still remains unexplained. An important clue to HRE may reside in the structural organization of hexanucleotide repeats as the GGGGCC sequence forms the non-canonical DNA structure called a G-quadruplex (G4) [13–17], and its complimentary sequence, GGCCCC, forms an intercalating motif (i-motif) [18, 19].

Here, we investigate the thermodynamic and kinetic properties of G4 configurations formed from the ALS/FTD-associated HRE. Recent work on other G4-forming sequences has indicated that the antiparallel configuration can be an intermediate in the process of folding to the more stable parallel G4 form, either existing transiently [20–22] or acting as a rapidly formed metastable state which slowly converts to a parallel G4 [23] over the course of minutes. We find the ALS/FTD-associated GGGGCC HRE, on the other hand, to be stable at room temperature indefinitely. Current computational and experimental evidence suggests that antiparallel G4s are generally separated from parallel G4s by a transition state barrier which leads to slow kinetics [23]. We experimentally manipulate this barrier through heating; if the transition state is enthalpically less stable and has greater entropy than the G4-forming states, heating should lead to a relative lowering of the free energy barrier, causing a "thermally activated" transition from the metastable antiparallel G4s to the stable parallel G4s.

* Correspondence to: dross@vcom.edu (D.R.); mohanty@hms.harvard.edu (V.M.); bmohanty@vcom.edu (B.K.M.).

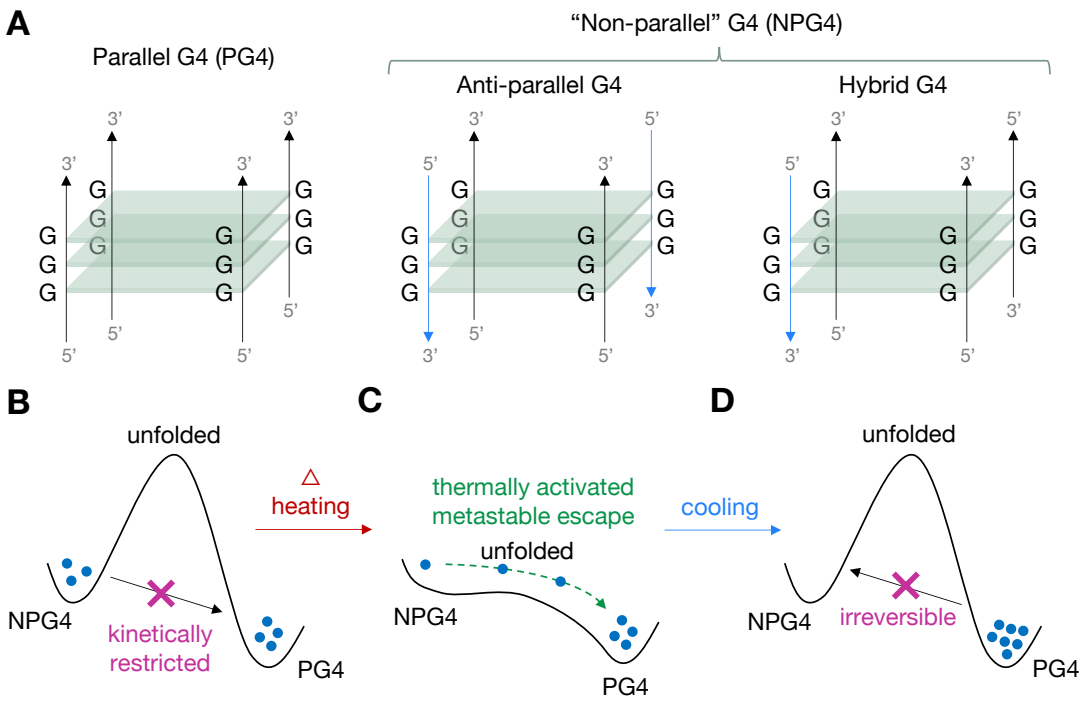


FIG. 1. **G-quadruplex structures and the proposed mechanism of thermally activated configurational homogenization.** (**A**) Three G4 configurations, with strand orientation determining whether a G4 is parallel, antiparallel, or hybrid. We group antiparallel and hybrid into a category called "non-parallel." (**B**) Initial heterogeneous population of G4s remain heterogeneous over long timescales due to transition state barrier separating NPG4s and PG4s. (**C**) A thermally activated metastable-to-stable transition: heating lowers the free energy barrier relative to the G4-forming states just enough to allow rapid transitions from the metastable NPG4 state to the stable PG4 state without unfolding G4s entirely. (**D**) Cooling the system once again imposes a transition state barrier for the reverse reaction for the now-homogeneous PG4 population.

In the present work, we combine theoretical and experimental approaches to probe the thermally activated metastable-to-stable transition of G4 configurations. First, we develop an analytical theory using a two-state model for G4 configuration prevalence. We then perform temperature-swept CD spectroscopy to obtain "3D CD" spectra [24, 25] showing how CD changes as a function of both wavelength and temperature. From these spectra, we observe the transition from a configurationally heterogeneous population of G4s into a homogeneous population of parallel G4s, which is direct evidence of the metastable-to-stable transition, in agreement with theory. Finally, we perform sequential temperature sweeps with a cooling period in between to demonstrate the irreversibility of the transition. Our work opens the door to understanding how the concentrations of different DNA G4 configurations may be stabilized or regulated; since G4s are known to facilitate intracellular RNA condensate formation in pathological disease states, these results carry implications for coupling between G4 configuration switches and disease-associated liquid-liquid phase separation of DNA [12, 26–28].

## II. RESULTS

### A. Two-state theory of thermally activated G4 homogenization

First, we show that a two-state analytical theory for the configurational state of G4s predicts a thermally activated metastable-to-stable transition which yields a homogeneous population of parallel G4s. We first introduce the theory before showing experimental results. We use $P$ to denote the parallel G4s (PG4s), and $N$ to denote the antiparallel or hybrid confirmations, which we informally group together as "non-parallel" G4s (NPG4s) since they are known to not be the thermodynamic ground state [23]; in Figure 1A we depict these different structures schematically. These form the two discrete states in our thermodynamic model. The double dagger ‡ superscript corresponds to the unfolded state, which we treat as an unstable transition state between the PG4 and NPG4 states.

To summarize the theory, we begin with a heuristic explanation. We expect that a heterogeneous population of NPG4s and PG4s are metastable for long times because of thermodynamic stability relative to the unfolded state, but they are separated by a tall transition state barrier

which leads to very slow kinetics (Figure 1B). Since the unfolded transition state has higher entropy than the G4 states, increasing the temperature lowers the relative height of the free energy barrier to transition from NPG4s to PG4s, making the metastable-to-stable transition suddenly possible even at short timescales, sending the population to primarily the PG4 state (Figure 1C). Cooling the system then traps the population homogeneously in the thermodynamically stable PG4 state (Figure 1D). Further heating would keep the population predominantly in the PG4 state due to its thermodynamic stability relative to NPG4.

Concretely, we define the Gibbs free energies, enthalpies, and entropies, working within the approximation that the formed G4s to have roughly equal enthalpy $H_0$ and approximating the transition state to have zero enthalpy:

$$\begin{aligned} G_P &= H_0 - T(S_N + \Delta S_{PN}) \\ G_N &= H_0 - TS_N \\ G^\ddagger &= -T(S_N + \Delta S_N^\ddagger). \end{aligned} \tag{1}$$

Here, $G$ values are Gibbs free energies, $T$ is the temperature, $S_N$ is the NPG4 entropy, $S_P$ is the PG4 entropy, $\Delta S_{PN} = S_P - S_N$, and $\Delta S_N^\ddagger = S^\ddagger - S_N$ is the entropy difference between the transition state and the NPG4 state. We take the parallel configuration entropy to be greater than the antiparallel entropy, which makes PG4 thermodynamically more stable than NPG4, consistent with ref. [23]. The system can undergo a chemical reaction:

$$N \underset{k_{NP}}{\overset{k_{PN}}{\rightleftharpoons}} P, \tag{2}$$

where the scaled reaction rates are Arrhenius-type:

$$k_{PN} = e^{-\frac{G^\ddagger - G_N}{RT}} = e^{\frac{H_0 + T\Delta S_N^\ddagger}{RT}} \tag{3}$$

and

$$k_{NP} = e^{-\frac{G^\ddagger - G_P}{RT}} = e^{\frac{H_0 + T\Delta S_N^\ddagger - T\Delta S_{NA}}{RT}} \tag{4}$$

assuming we have divided the reaction rates by the empirical transition attempt frequency $A_0$ and $R$ is the ideal gas constant.

We let $p(t;T)$ be the fraction of PG4s at time $t$ and fixed temperature $T$, so $1 - p(t;T)$ is the fraction of NPG4s. We can write the rate equation

$$\frac{1}{A_0}\frac{\mathrm{d}p}{\mathrm{d}t} = -k_{NP}p(t;T) + k_{PN}(1 - p(t;T)). \tag{5}$$

For mathematical convenience, we define a rescaled time $\tau = A_0 t$ which now defines the timescale based on the empirical transition attempt frequency. The solution to the above differential equation is

$$\begin{aligned} p(\tau;T) = e^{-(k_{NP}+k_{PN})\tau}\left(p(0;T) - \frac{k_{PN}}{k_{PN}+k_{NP}}\right) \\ + \frac{k_{PN}}{k_{PN}+k_{NP}}, \end{aligned} \tag{6}$$

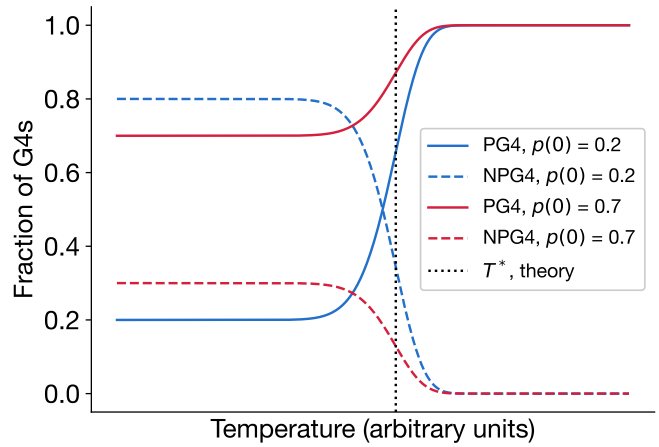


FIG. 2. **Theoretical fraction of PG4s and NPG4s as a function of temperature indicates phase transition-like behavior from a starting PG4 fraction $p(0;T)$ and NPG4 fraction $1 - p(0;T)$ to a (nearly) homogeneous population of PG4s at higher temperatures.** The homogenization temperature $T^*$, approximated by the analytical expression in eq. (8), is plotted (black, dotted). Two starting fractions, $p(0;T) = 0.2$ (blue) and $p(0;T) = 0.7$ (red) are plotted.

which can be expressed in terms of $R$, $T$, $H_0$, $\Delta S_N^\ddagger$, $\Delta S_{NA}$, and $\tau$.

For some initial fraction $p(0;T)$ of PG4s, eq. (6) gives the fraction $p(\tau;T)$ of PG4s after some fixed scaled time $\tau$. Plotting $p(\tau;T)$ versus $T$ shows that, independent of $p(0;T)$, there is some temperature $T^*$ around which a heterogeneous starting population becomes nearly all PG4s (Figure 2). This is evidence of a *thermally activated* transition of the metastable NPG4s to the stable PG4s, and represents complete or almost complete homogenization of the G4 population to the parallel configuration. This is not a true thermodynamic phase transition because the ground state configuration does not change; but rather, the theory predicts a kinetic phenomenon where a long-lived metastable state transitions to a stable one when the relative height of the transition state barrier is reduced via heating.

Using first-order perturbation theory, we obtain an approximate analytical expression for the inflection point of the sigmoidal transition curve, which we call the homogenization temperature $T^*$ (see Supplementary Note 1 for derivation):

$$T^* \approx T_0 - \frac{2R^2T_0^3}{H_0^2 + 2R^2T_0^2}, \tag{7}$$

where $T_0$ is the zeroth-order solution

$$T_0 = \frac{H_0}{\Delta S_{PN} - \Delta S_N^\ddagger - R\log\tau - R\log\left(1 + e^{\frac{\Delta S_{PN}}{R}}\right)}. \tag{8}$$

We plot the analytical approximation to the transition temperature in Figure 2. The existence of such an inflec-

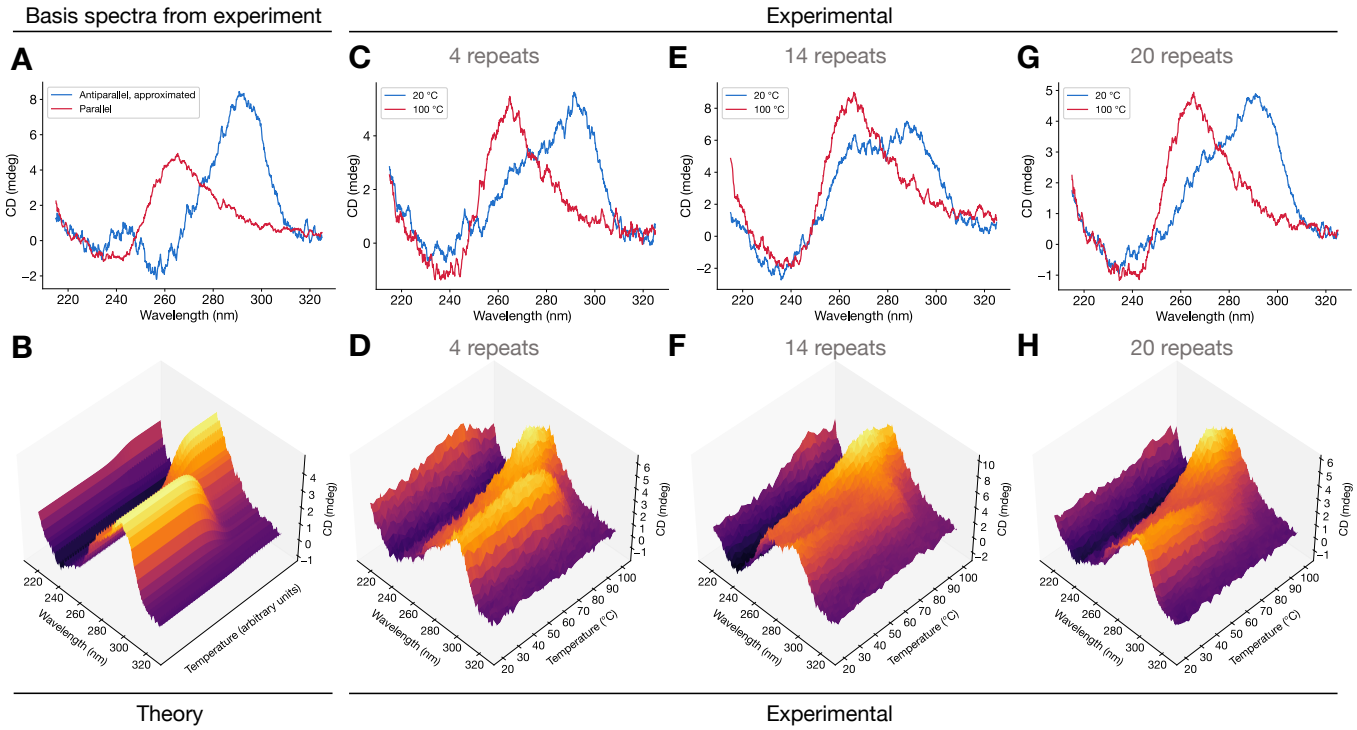


FIG. 3. **3D CD spectra reveal thermally activated metastable-to-stable transition from NPG4s to PG4s**. (**A**) Experimentally determined PG4 basis CD spectrum $\epsilon_P(\lambda)$ (red) and approximated antiparallel basis CD spectrum $\epsilon_A(\lambda)$ (blue), used to construct theoretical 3D CD spectrum. (**B**) Theoretical 3D CD spectrum computed from eq. (9) by using eq. (6) and basis spectra from (A), plotted on an arbitrary temperature axis, reveals a CD positive peak shift from near 290 nm to near 265 nm, which is an indicator of NPG4s converting to PG4s. (**C**,**E**,**G**) CD spectra at 20 °C (blue) and 70 °C (red) for 4, 14, and 20 repeats, respectively. (**D**,**F**,**H**) 3D CD spectra showing temperature swept circular dichroism, experimentally revealing the metastable-to-stable transition from NPG4s to PG4s as predicted by theory.

tion point is mathematically indicative of the thermally activated phase transition-like behavior, which we examine experimentally through CD spectroscopy.

### B. Experimental CD spectra reveal thermally activated metastable-to-stable transition

To probe the configurational composition of G4 populations experimentally, we conducted temperature-swept CD spectroscopy of DNA oligonucleotides containing the GGGGCC hexanucleotide repeat. This involved taking CD spectra of the same DNA sample as the temperature was incrementally increased from 20 °C to 100 °C in increments of 2 °C, with 30 second wait periods before CD measurement. The resulting data were 3D CD spectra which visualize circular dichroism as a surface plot against both wavelength and temperature.

Before introducing the experimental results, we first discuss the standard spectroscopic signatures of various G4 conformations, and the 3D CD prediction from theory. Parallel, antiparallel, and hybrid G4s have distinct CD peaks. Parallel G4s have a negative peak at 240 nm and a larger positive peak around 260 nm, antiparallel G4s have a negative peak at 260 nm and a positive peak at 295 nm, and hybrid G4s show positive peaks at both 260 nm and 295 nm and a negative peak around 240 nm [29]. Based on these spectroscopic signatures, we posit that the hybrid conformation may be difficult to distinguish from a mixture of antiparallel, hybrid, and parallel G4s, so for simplicity we take the parallel and antiparallel structures to be the fundamental "basis spectra" for our theoretical analysis.

CD spectra matching the features of parallel and antiparallel G4s are plotted in Figure 3A. The parallel G4 spectrum, obtained at 100 °C, closely matches published spectra [29], and the antiparallel CD spectrum is approximated by subtracting off the parallel contribution to an otherwise mixed or hybrid-appearing spectrum at 20 °C (see Methods for details). These parallel and antiparallel spectra (Figure 3A), $\epsilon_P(\lambda)$ and $\epsilon_A(\lambda)$, are functions of wavelength $\lambda$ and are used to obtain theoretical 3D CD spectra.

Treating the 3D CD spectrum [24, 25] as a linear combination of the basis spectra weighted by the PG4 and NPG4 fractions from eq. (6), we have an expression for the theoretical 3D CD spectrum

$$\epsilon_{3\mathrm{D}}(\lambda, T) = p(\tau; T)\epsilon_P(\lambda) + (1 - p(\tau; T))\epsilon_A(\lambda), \qquad (9)$$

which is plotted in Figure 3B. Notably, the homogeniza-

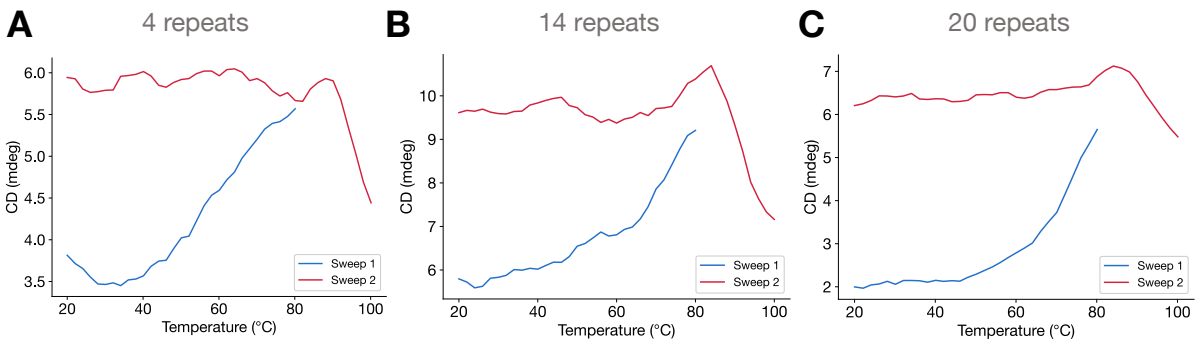


FIG. 4. **Thermally activated metastable-to-stable G4 transition is irreversible**. (**A**) For 4 hexanucleotide repeats, the first temperature sweep (blue) shows a CD increase at 269 nm as temperature increases from 20 °C up to 80 °C, indicative of a transition to PG4s. After cooling back to 20 °C, the second temperature sweep (red) shows stability of the CD peak from 20 °C up to roughly 80 °C, after which CD decreases, either due to G4 melting (or possibly NPG4 formation). (**B**) Same as (A), but with 14 hexanucleotide repeats with CD measured at 264 nm. (**C**) Same as (A) and (B), but with 20 hexanucleotide repeats with CD measured at 267 nm.

tion transition predicted by theory in Figure 2 manifests as a shift in the positive peak around 290 nm to a peak around 260 nm, as the temperature increases past the transition temperature $T^*$.

Experimental 3D CD spectra of hexanucleotide repeats with various copy numbers closely match the theoretical finding. For 4 repeats, 14 repeats, and 20 repeats, we observe that the initial samples at 20 °C show peaks that match hybrid or a mixed population of antiparallel, hybrid, and parallel G4s—an indication that we likely have a configurationally heterogeneous population of G4s (Figure 3C/E/G). The final CD spectra, at the maximum temperature of 100 °C, all match the spectroscopic signature of parallel G4s, suggesting homogenization at some intermediate temperature (Figure 3C/E/G). The 3D CD spectra showing the continuous dependence on both wavelength and temperature indicate, that regardless of copy number, the 290 nm peak is stable from the starting temperature 20 °C up to roughly 70 °C, after which the 260 nm peak rapidly rises to prominence at higher temperatures (Figure 3D/F/H). The profile of the experimental 3D CD spectra match the theoretical spectrum (Figure 3B) qualitatively, supporting the notion of a thermally activated metastable-to-stable transition. We performed the same temperature-swept CD spectroscopy for additional copy number variants and observed consistent results (Figure S1). Moreover, to confirm that CD spectroscopic changes associated with the transition is not limited to oligonucleotides from one commercial source, we tested the effect of temperature increase on CD with oligonucleotides from another commercial source and obtained the same results.

### C. Irreversibility of G4 homogenization

As explained by our theory, the metastable-to-stable transition should be irreversible. The initial kinetic trap which stabilizes NPG4s over long timescales disappears with thermal activation, and once the population converts to PG4s, the reverse reaction is unfavorable thermodynamically. As a result, a homogeneous PG4 population should remain homogeneous, even if the system is cooled back to room temperature after the initial temperature sweep. A second temperature sweep should thus lead to little change in the configurational composition, which means the CD spectrum should also not change after the first sweep. This would be a key indication of irreversibility.

To investigate irreversibility, we sequentially performed a temperature sweep followed by a cooling period back to near room temperature followed by a second temperature sweep. For oligonucleotide samples with various copy numbers, we measured CD at a fixed wavelength corresponding to the principle positive peak for PG4s during each of the two temperature sweeps.

In Figure 4A-C, we show for 4, 14, and 20 hexanucleotide repeats that circular dichroism around 264-269 nm generally increases with temperature during the sweep from 20 °C up to 80 °C, consistent with previous section's experimental results indicating a predominant presence of PG4s and a decline in NPG4s. After cooling back to 20 °C, during the second sweep, we observed that even at 20 °C, the CD magnitude remained near the value which it had reached at 80 °C during the first sweep. This is a clear signature of irreversibility; after the first sweep homogenized the G4 population to the parallel configuration, cooling did not lead to a conversion back to the initial heterogeneous distribution. Subsequently, throughout the second sweep, from 20 °C up to 80 °C, CD magnitude remained roughly constant, consistent with the theory that the homogeneous PG4 population remained homogeneous during the second sweep due to the thermodynamic stability of the PG4 configuration. Above 80 °C, we observed sharp declines in the CD, indicating that either the transition state barrier was

now low enough once again to permit the reverse reaction back to NPG4 states, or that we exceeded the melting temperature of the PG4 configuration for this particular hexanucleotide repeat, leading to unfolding of PG4s into linear strands. The same irreversibility experiments for additional copy number variants demonstrated consistent results (Figure S1), supporting the theory of irreversible thermally activated homogenization via a metastable-to-stable transition.

## III. DISCUSSION

In this work, we have explored the thermodynamic and kinetic relationships between parallel and non-parallel G4 configurations of the GGGGCC hexanucleotide repeat expansion which appears in ALS/FTD. We provided an analytical two-state theory to explain the thermally activated transition from the metastable NPG4s to a stable PG4s. By performing temperature-swept CD measurements, we plotted 3D CD spectra which showed a clear change in CD spectral shape indicative of homogenization of G4s to the parallel confriguration above a threshold temperature. We then demonstrated experimentally that this transition was indeed irreversible, as expected from a metastable-to-stable transition, by performing sequential temperature sweeps with a cooling step in between.

These results indicate that regulation of the transition state barrier separating PG4s and NPG4s can lead to substantial changes in the composition of the G4 population. Since G4s are present *in vivo* [30], we hypothesize that there may be intracellular enzymatic regulation of the free energy barrier separating the configurations, since thermally activated transitions observed here require temperatures that are too high for human life.

Moreover, our work opens the door to further investigation regarding how G4 configurations may be coupled to biomolecular condensate formation observed in pathological cell states. An understanding of how NPG4-PG4 transitions, unimolecular G4 versus multimolecular G4 formation, and liquid-liquid phase separation are related to each other is still an open question which will be addressed in our forthcoming work.

## IV. METHODS

*DNA oligonucleotides.* DNA oligonucleotides used in these studies were obtained commercially from Millipore (Burlington, MA) and Eurofins. All oligonucleotides were purified by high-performance liquid chromatography, shipped dry, and reconstituted with TE (10 mM Tris and 1 mM EDTA) or nuclease free water (Invitrogen) to 100 pmol/$\mu$L final concentration.

*Buffers.* The buffers used in these studies are 10 mM sodium cacodylate and 100 mM KCl (NaCac) at pH 7.4.

*CD measurements.* All CD measurements were carried out with a Jasco J-1500 CD spectrophotometer with a single-position Peltier-thermo cell holder. Samples were measured in 1 mm path length Starna cuvettes. For the 3D CD experiments, temperature sweeps were set to start at 20 °C and end at 100 °C, with a temperature gradient of 1 °C/min and a wait time of 30 seconds between measurements, which occurred at intervals of 2 °C. The temperature gradient was paused during measurement. The starting condition was set to maintaining the initial temperature within ±0.1 °C for 10 seconds. For the irreversibility experiments, the first sweep was conducted from 20 °C to 80 °C, while the second sweep was conducted from 20 °C to 100 °C. For 4, 14, and 20 repeats, the observation wavelength was set to 269 nm, 264 nm, and 267 nm, respectively. For the irreversibility experiment plots, CD lines were smoothed with a moving average filter using a convolution kernel with a 5-datapoint width.

*Basis spectra and theoretical 3D CD spectrum.* Because a pure antiparallel G4 sample was not available, the approximated basis spectrum was obtained by assuming an equal mixture of parallel and antiparallel G4s in the spectrum obtained at 20 °C during the 3D CD experiment using 20 repeats. The spectrum obtained at 100 °C was assumed to be all PG4s and was assigned as the basis spectrum for PG4s. The approximated antiparallel spectrum was obtained by subtracting off the basis PG4 spectrum weighted by 0.5, from the 20 °C, and finally dividing by 0.5.

## V. ACKNOWLEDGEMENTS

*Funding and resource acknowledgement.* This work was supported by VCOM's REAP grants 1032453 and 1302559 (to B.K.M.), a Hertz Foundation Fellowship (to V.M.), and a PD Soros Fellowship (to V.M.). The authors declare no known conflict of interest.

*Code and data availability.* Code and data will be made publicly available upon publication.

## VI. AUTHOR CONTRIBUTIONS

D.R., D.E., V.M., and B.K.M. developed the project idea. D.R., O.L., O.M., S.B., S.D., R.B., R.D., and B.K.M. conducted experiments. V.M. contributed analytical theory. D.R., V.M., and B.K.M. analyzed the data. D.R., V.M., and B.K.M. wrote the manuscript. B.K.M. supervised the project.

[1] E. L. Feldman, S. A. Goutman, S. Petri, L. Mazzini, M. G. Savelieff, P. J. Shaw, and G. Sobue, Amyotrophic lateral sclerosis, *The Lancet* **400** (2022) 1363–1380.

[2] S. Zarei, K. Carr, L. Reiley, K. Diaz, O. Guerra, P. F. Altamirano, W. Pagani, D. Lodin, G. Orozco, and A. Chinea, A comprehensive review of amyotrophic lateral sclerosis, *Surgical Neurology International* **6** (2015) 171.

[3] R. J. Mead, N. Shan, H. J. Reiser, F. Marshall, and P. J. Shaw, Amyotrophic lateral sclerosis: a neurodegenerative disorder poised for successful therapeutic translation, *Nature Reviews Drug Discovery* **22** (2023) 185–212.

[4] J. R. Hodges and B. Miller, The classification, genetics and neuropathology of frontotemporal dementia. Introduction to the special topic papers: Part I, *Neurocase* **7** (2001), no. 1 31–35.

[5] J. R. Burrell, G. M. Halliday, J. J. Kril, L. M. Ittner, J. Götz, M. C. Kiernan, and J. R. Hodges, The frontotemporal dementia-motor neuron disease continuum, *The Lancet* **388** (2016) 919–931.

[6] S. Ghosh and C. F. Lippa, Clinical Subtypes of Frontotemporal Dementia, *American Journal of Alzheimer's Disease and Other Dementias* **30** (2015) 653–661.

[7] L. Tzeplaeff, S. Wilfling, M. V. Requardt, and M. Herdick, Current State and Future Directions in the Therapy of ALS, *Cells* **12** (2023) 1523.

[8] R. Rademakers, C9ORF72 repeat expansions in ALS and FTD patients worldwide, *Lancet neurology* **11** (2012) 297–298.

[9] A. E. Renton, E. Majounie, A. Waite, J. Simón-Sánchez, S. Rollinson, J. R. Gibbs, J. C. Schymick, H. Laaksovirta, J. C. van Swieten, L. Myllykangas, H. Kalimo, A. Paetau, Y. Abramzon, A. M. Remes, A. Kaganovich, S. W. Scholz, J. Duckworth, J. Ding, D. W. Harmer, D. G. Hernandez, J. O. Johnson, K. Mok, M. Ryten, D. Trabzuni, R. J. Guerreiro, R. W. Orrell, J. Neal, A. Murray, J. Pearson, I. E. Jansen, D. Sondervan, H. Seelaar, D. Blake, K. Young, N. Halliwell, J. B. Callister, G. Toulson, A. Richardson, A. Gerhard, J. Snowden, D. Mann, D. Neary, M. A. Nalls, T. Peuralinna, L. Jansson, V.-M. Isoviita, A.-L. Kaivorinne, M. Hölttä-Vuori, E. Ikonen, R. Sulkava, M. Benatar, J. Wuu, A. Chiò, G. Restagno, G. Borghero, M. Sabatelli, D. Heckerman, E. Rogaeva, L. Zinman, J. D. Rothstein, M. Sendtner, C. Drepper, E. E. Eichler, C. Alkan, Z. Abdullaev, S. D. Pack, A. Dutra, E. Pak, J. Hardy, A. Singleton, N. M. Williams, P. Heutink, S. Pickering-Brown, H. R. Morris, P. J. Tienari, and B. J. Traynor, A Hexanucleotide Repeat Expansion in *C9ORF72* Is the Cause of Chromosome 9p21-Linked ALS-FTD, *Neuron* **72** (2011) 257–268.

[10] Y. Sonobe, S. Lee, G. Krishnan, Y. Gu, D. Y. Kwon, F.-B. Gao, R. P. Roos, and P. Kratsios, Translation of dipeptide repeat proteins in C9ORF72 ALS/FTD through unique and redundant AUG initiation codons, *eLife* **12** (2023) e83189.

[11] S. Wang, M. J. Latallo, Z. Zhang, B. Huang, D. G. Bobrovnikov, D. Dong, N. M. Livingston, W. Tjoeng, L. R. Hayes, J. D. Rothstein, L. W. Ostrow, B. Wu, and S. Sun, Nuclear export and translation of circular repeat-containing intronic RNA in C9ORF72-ALS/FTD, *Nature Communications* **12** (2021) 4908.

[12] F. Raguseo, Y. Wang, J. Li, M. Petrić Howe, R. Balendra, A. Huyghebaert, D. M. Vadukul, D. A. Tanase, T. E. Maher, L. Malouf, R. Rubio-Sánchez, F. A. Aprile, Y. Elani, R. Patani, L. Di Michele, and M. Di Antonio, The ALS/FTD-related C9orf72 hexanucleotide repeat expansion forms RNA condensates through multimolecular G-quadruplexes, *Nature Communications* **14** (2023) 8272.

[13] D. Sen and W. Gilbert, Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis, *Nature* **334** (1988) 364–366.

[14] K. Reddy, B. Zamiri, S. Y. R. Stanley, R. B. Macgregor, and C. E. Pearson, The Disease-associated r(GGGGCC)n Repeat from the C9orf72 Gene Forms Tract Length-dependent Uni- and Multimolecular RNA G-quadruplex Structures, *The Journal of Biological Chemistry* **288** (2013) 9860–9866.

[15] E. G. Conlon, L. Lu, A. Sharma, T. Yamazaki, T. Tang, N. A. Shneider, and J. L. Manley, The C9ORF72 GGGGCC expansion forms RNA G-quadruplex inclusions and sequesters hnRNP H to disrupt splicing in ALS brains, *eLife* **5** (2016) e17820.

[16] B. Zhou, Y. Geng, C. Liu, H. Miao, Y. Ren, N. Xu, X. Shi, Y. You, T. Lee, and G. Zhu, Characterizations of distinct parallel and antiparallel G-quadruplexes formed by two-repeat ALS and FTD related GGGGCC sequence, *Scientific Reports* **8** (2018) 2366.

[17] Y. Geng, C. Liu, Q. Cai, Z. Luo, H. Miao, X. Shi, N. Xu, C. P. Fung, T. T. Choy, B. Yan, N. Li, P. Qian, B. Zhou, and G. Zhu, Crystal structure of parallel G-quadruplex formed by the two-repeat ALS- and FTD-related GGGGCC sequence, *Nucleic Acids Research* **49** (2021) 5881–5890.

[18] H. Abou Assi, M. Garavís, C. González, and M. J. Damha, i-Motif DNA: structural features and significance to cell biology, *Nucleic Acids Research* **46** (2018) 8038–8056.

[19] K. Gehring, J.-L. Leroy, and M. Guéron, A tetrameric DNA structure with protonated cytosine-cytosine base pairs, *Nature* **363** (1993) 561–565.

[20] R. D. Gray, J. O. Trent, and J. B. Chaires, Folding and Unfolding Pathways of the Human Telomeric G-Quadruplex, *Journal of Molecular Biology* **426** (2014) 1629–1650.

[21] M.-L. Zhang, Y.-P. Xu, A. Kumar, Y. Zhang, and W.-Q. Wu, Studying the Potassium-Induced G-Quadruplex DNA Folding Process Using Microscale Thermophoresis, *Biochemistry* **58** (2019) 3955–3959.

[22] A. Marchand and V. Gabelica, Folding and misfolding pathways of G-quadruplex DNA, *Nucleic Acids Research* **44** (2016) 10999–11012.

[23] J. Xia, J. Chen, J. Zhou, M. Cheng, X. Zhuang, C. Cai, H. Ju, J.-L. Mergny, and J. Zhou, Antiparallel G-Quadruplex Formation Hinders Conversion to a Parallel Topology, *The Journal of Physical Chemistry B* **128** (2024) 11077–11087.

[24] M. M. Dailey, M. C. Miller, P. J. Bates, A. N. Lane, and J. O. Trent, Resolution and characterization of the structural polymorphism of a single quadruplex-forming sequence, *Nucleic Acids Research* **38** (2010) 4877–4888.

[25] A. Nicosia, F. Vento, G. Marletta, G. M. L. Messina, C. Satriano, V. Villari, N. Micali, M. T. De Martino, M. J. G. Schotman, and P. G. Mineo, Porphyrin-Based Supramolecular Flags in the Thermal Gradients' Wind: What Breaks the Symmetry, How and Why, *Nanomaterials* **11** (2021) 1673.

[26] W. Wang, D. Li, Q. Xu, J. Cheng, Z. Yu, G. Li, S. Qiao, J. Pan, H. Wang, J. Shi, T. Zheng, and G. Sui, G-quadruplexes promote the motility in MAZ phase-separated condensates to activate CCND1 expression and

contribute to hepatocarcinogenesis, *Nature Communications* **15** (2024) 1045.

[27] M. Tsuruta, T. Torii, K. Kohata, K. Kawauchi, H. Tateishi-Karimata, N. Sugimoto, and D. Miyoshi, Controlling liquid–liquid phase separation of G-quadruplex-forming RNAs in a sequence-specific manner, *Chemical Communications* **58** (2022) 12931–12934.

[28] Z. Gao, J. Yuan, X. He, H. Wang, and Y. Wang, Phase Separation Modulates the Formation and Stabilities of DNA Guanine Quadruplex, *JACS Au* **3** (2023) 1650–1657.

[29] M. Moriya, T. Oyama, M. Goto, K. Ikebukuro, and W. Yoshida, Protocol for circular dichroism spectral analysis of the thermal stability of CpG-methylated quadruplex structures, *STAR Protocols* **6** (2025) 103646.

[30] Z.-H. Zhang, S. H. Qian, D. Wei, and Z.-X. Chen, In vivo dynamics and regulation of DNA G-quadruplex structures in mammals, *Cell & Bioscience* **13** (2023) 117.

# Supplementary Materials for "Thermally activated irreversible homogenization of G-quadruplexes in an ALS/FTD-associated nucleotide expansion"

Daniel Ross, Olivia Lewis, Olivia McLean, Sundeep Bhanot, Shane Donahue, Rachael Baker, Randi Dias, David Eagerton, Vaibhav Mohanty, and Bidyut K. Mohanty

### Supplementary Note 1: Perturbative approximation of homogenization temperature

We estimate the homogenization temperature as the inflection point of the curves in Figure 2, which is the $T^*$ for which

$$\left.\frac{\partial^2 p(\tau;T)}{\partial T^2}\right|_{T=T^*} = 0, \tag{S1}$$

which, after dividing by prefactors that can never be zero or do not depend on $T$, gives us

$$f(T^*) \equiv H_0\tau e^{\frac{H_0+T^*(\Delta S_N^\ddagger+\Delta S_{PN})}{RT^*}} + H_0\tau e^{\frac{H_0+\Delta S_N^\ddagger T^*}{RT^*}} - e^{\frac{\Delta S_{PN}}{R}}(H_0+2RT^*) = 0. \tag{S2}$$

In the limit of small $T$, we call the zeroth-order solution $T_0$ and can write

$$H_0\tau e^{\frac{H_0}{RT_0}+\frac{\Delta S_N^\ddagger+\Delta S_{PN}}{R}} + H_0\tau e^{\frac{H_0}{R}+\frac{\Delta S_N^\ddagger}{RT_0}} - e^{\frac{\Delta S_{PN}}{R}} H_0 = 0, \tag{S3}$$

which can be solved to obtain the zeroth-order solution for the transition temperature

$$T_0 = \frac{H_0}{\Delta S_{PN} - \Delta S_N^\ddagger - R\log\tau - R\log\left(1+e^{\frac{\Delta S_{PN}}{R}}\right)}. \tag{S4}$$

To obtain a first-order perturbative correction, we add a small correction to the zeroth-order approximation:

$$T^* \approx T_0 + \delta T, \tag{S5}$$

and we have

$$0 = f(T^*) = f(T_0+\delta T) \approx f(T_0) + \delta T \left.\frac{\partial f}{\partial T}\right|_{T=T_0}, \tag{S6}$$

which gives us

$$\delta T = -\frac{f(T_0)}{\left.\frac{\partial f}{\partial T}\right|_{T=T_0}}. \tag{S7}$$

First, we use eq. (S3) to rewrite

$$f(T_0) = -2RT_0 e^{\frac{\Delta S_{PN}}{R}}. \tag{S8}$$

We now calculate the derivative

$$\left.\frac{\partial f}{\partial T}\right|_{T=T_0} = -\frac{H_0^2\tau\left(e^{\frac{H_0+T_0(\Delta S_N^\ddagger+\Delta S_{PN})}{RT_0}} + e^{\frac{H_0+\Delta S_N^\ddagger T_0}{RT_0}}\right) + 2R^2T_0^2 e^{\frac{\Delta S_{PN}}{R}}}{RT_0^2}. \tag{S9}$$

Using eq. (S3), we can write

$$\left.\frac{\partial f}{\partial T}\right|_{T=T_0} = -e^{\frac{\Delta S_{PN}}{R}}\frac{H_0^2+2R^2T_0^2}{RT_0^2}. \tag{S10}$$

We can now express the perturbative correction as

$$\delta T = -\frac{-2RT_0 e^{\frac{\Delta S_{PN}}{R}}}{-e^{\frac{\Delta S_{PN}}{R}}\frac{H_0^2+2R^2T_0^2}{RT_0^2}} = -\frac{2R^2T_0^3}{H_0^2+2R^2T_0^2}, \tag{S11}$$

giving the perturbative solution

$$T^* \approx T_0 - \frac{2R^2T_0^3}{H_0^2+2R^2T_0^2}. \tag{S12}$$

## Supplementary Figure

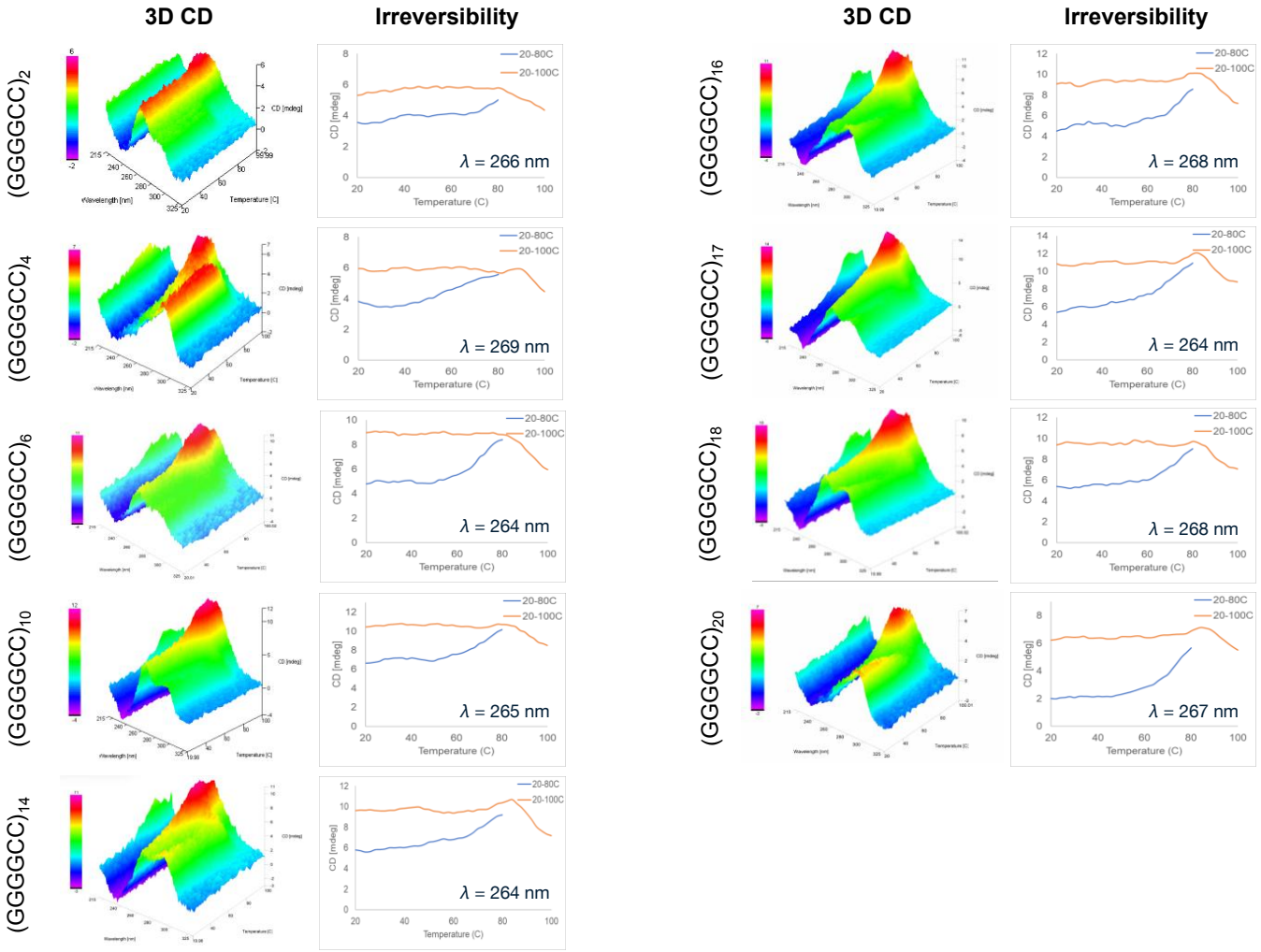


FIG. S1. **3D CD spectra and irreversibility data for additional oligonucleotides with varying hexanucleotide repeat copy numbers.** 3D CD spectra demonstrate the metastable-to-stable homogenization transition consistent with theory, and consistent across hexanucleotide repeat copy numbers. Irreversibility data show stability of the CD peak during the second temperature sweep (20 °C to 100 °C, orange) after it was reached at the end of the first temperature sweep (20 °C to 80 °C, blue). Irreversibility data are consistent across hexanucleotide repeat copy numbers. The wavelengths used to capture irreversibility data for each copy number are listed and are all in the 264-269 nm.