Supporting Information

Understanding Self-Assembled Pseudoisocyanine Dye Aggregates in DNA Nanostructures and their Exciton Relay Transfer Capabilities

Matthew Chiriboga[1,2], Sebastian A. Diaz[1*], Divita Mathur[1,3], David A. Hastman[1,4], Joseph S. Melinger[5], Remi Veneziano[2], and Igor L. Medintz[1*]

[1]Center for Bio/Molecular Science & Engineering Code 6900
[5]Electronics Science and Technology Division Code 6800
U.S. Naval Research Laboratory, 4555 Overlook Ave. S.W., Washington, DC 20375, USA

[2]Volgenau School of Engineering, Department of Bioengineering, Institute for Advanced Biomedical Research, George Mason University, Manassas, VA 22030, USA

[3]College of Science, George Mason University, Fairfax, VA 22030, USA

[4]A. James Clark School of Engineering, Fischell Department of Bioengineering, University of Maryland College Park, College Park, MD 20742, USA

*Corresponding Authors
Email: Sebastian.Diaz@nrl.navy.mil , Igor.Medintz@nrl.navy.mil

Contents:

Supplemental Methods

FRET Calculations

FRET was analyzed in accordance with Förster theory[1-4] where the rate of energy transfer is given by equation (S1):

$$k_T(r) = \frac{Q_D k^2}{\tau_D r^6}\left(\frac{9000(\ln 10)}{128\pi^5 N n^4}\right)\int_0^\infty F_D(\lambda)\varepsilon_A(\lambda)\lambda^4\, d\lambda \qquad \text{Eq. (S1)}$$

where Q_D is the quantum yield of the donor in the absence of acceptor, n is the refractive index (1.333 for aqueous buffer), N is Avogadro's number, r is the distance separating the donor and τ_D is the fluorescence lifetime of the donor in the absence of the acceptor. The k^2 value refers to the orientation of the donor and acceptor transition dipoles, which is assumed to be $\frac{2}{3}$ for random dynamic averaging of dyes bound by hydrocarbon linkers in aqueous buffer. $F_D(\lambda)$ is the fluorescence intensity of the donor at a given wavelength (nm) normalized to the total area under the spectral curve. $\varepsilon_A(\lambda)$ is the extinction coefficient of the acceptor at a given wavelength in units of cm^{-1} M^{-1}. Rearranging the equation, it can be written in terms of the Förster distance (R_0) as seen in equation (S2).

$$R_0^6 = \frac{9000(\ln 10)Q_D k^2}{128\pi^5 N n^4}\int_0^\infty F_D(\lambda)\varepsilon_A(\lambda)\lambda^4\, d\lambda \qquad \text{Eq. (S2)}$$

Using R_0, FRET efficiencies (E) at any given separation distance r can be predicted.

$$E(r) = \frac{R_0^6}{R_0^6 + r^6} \qquad \text{Eq. (S3)}$$

Equation (S3) illustrates how the efficiency of energy transfer decays by r^6.

Spectral Decomposition

Fluorescence emission spectra were collected from 420 nm to 800 nm at 20°C with an n of 3 and the presented spectra are the average of the three repeats. The excitation wavelength was set to 395 nm with a slit width of 10 nm. After the spectra were collected they were imported into A|E UV-Vis-IR Spectral Software, which is a freely available spectral viewer written in MATLAB from FLuorTools.com.[5] This software allows for the manipulation of spectra such as baseline subtractions and decomposition of linearly combined dye component spectral peaks. This allows the user to solve for the relative contribution of individual dye components when the emission ranges overlap such as for PIC and A_{647}. Therefore, the experimentally measured combined spectra can be represented as:

$$B = x_{405}A_{405} + x_{PIC}A_{PIC} + x_{647}A_{647} \qquad \text{Eq. (S4)}$$

Here, x_i represents the unknown fraction of each dye contribution and A_i represents the isolated dye component i. In order to account for differences in A_{647} peak shape and shifting due to PIC interactions, an experimentally derived dye component was used to fit the measured spectra. In order to obtain the isolated A_{647} component for each structure, the unlabeled-PIC (-D_{405}, +J-bit, -A_{647}) control was subtracted from the D_{405}-PIC-A_{647} (+D_{405}, +J-bit, +A_{647}) system for all scaffolds.

Each of these isolated A_{647} components were then averaged to make a combined component, which was subsequently used to fit each measured spectra. The resulting fits and component contributions can be seen in **(Figure S12-13).**

Quantum Yield Measurements

Fluorescence Quantum yields of D_{405} and A_{647} were measured using a 0.3 cm path length measurement cell. Each sample had excitation and emission slit widths of 3 nm. All samples had a measured absorbance of less than 0.1 at the excitation wavelength. Dye relative quantum yield, Φ_s, was calculated following the method in Würth et al.[6]

$$\Phi_s = \Phi_{st} \times \frac{I_s}{I_{st}} \times \frac{f_{st}}{f_s} \times \left(\frac{n_s}{n_{st}}\right)^2 \qquad \text{Eq. (S5)}$$

Where I refers to the integrated fluorescence emission spectrum starting at 375 nm for D_{405} and 610 for A_{647} and ending at 750 nm for all measurements, f is the fraction of light absorbed ($1 - 10^{-A}$) with A being the absorbance at the respective excitation wavelengths of 350 nm for D_{405} and 600 nm for A_{647}, and n is the refractive index of the measurement medium. The subscripts s and st refer to "sample" and 'standard" respectively. D_{405} was measured relative to a quinine bisulfate standard ($\Phi = 0.502$)[7] in 0.1 N of sulfuric acid ($n = 1.33$). A_{647} was measured relative to oxazine720 ($\Phi = 0.63$)[8] in methanol ($n = 1.32$).

Melting Temperature Calculations

Using Python simple melting temperatures (T_m) as well as melting temperatures adjusted to 10 mM sodium (T_{ms}) were approximated based on strand sequence both assuming 50 µM primer and pH 7.0. Simple melting temperatures were calculated using equation (S6):[9]

$$T_m = 64.9 + 41 \times \frac{y+z-16.4}{w+x+y+z} \qquad \text{Eq. (S6)}$$

and salt adjusted melting temperatures were calculated following equation (S7):[9]

$$T_{ms} = 100.5 + 41 \times \frac{y+z}{w+x+y+z} - \left(\frac{820}{w+x+y+z}\right) + 16.6 \times \log_{10}([Na^+]) \qquad \text{Eq. (S7)}$$

where w, x, y, and z correspond to the number of A, T, C and G bases respectively and $[Na^+]$ corresponds to the Molar concentration of sodium ions *i.e.* 0.01 M.

Melting Temperature Measurements

Circular dichroism measurements at 250 nm (CD250) and absorbance measurements at 260 nm (A260) were collected using a J-1500 JASCO spectrophotometer as the average of 2 independent measurements. Using a 1 cm path cuvette, each sample contained 2.5 µM DX-tile in HEPES-Mg buffer. For the GC and AT10 samples the sample temperature was sequentially increased from 10°C to 90°C with measurements taken after each 10°C increment change. For the AT30 sample the temperature was sequentially increased from 10°C to 80°C with measurements taken after each 2°C increment changes. Both the CD250 and the A260 data were fit using the origin software

Boltzmann fitting function and the resulting fit X_0 value used as a measured melting temperature value as reported in F**igure S10E**.

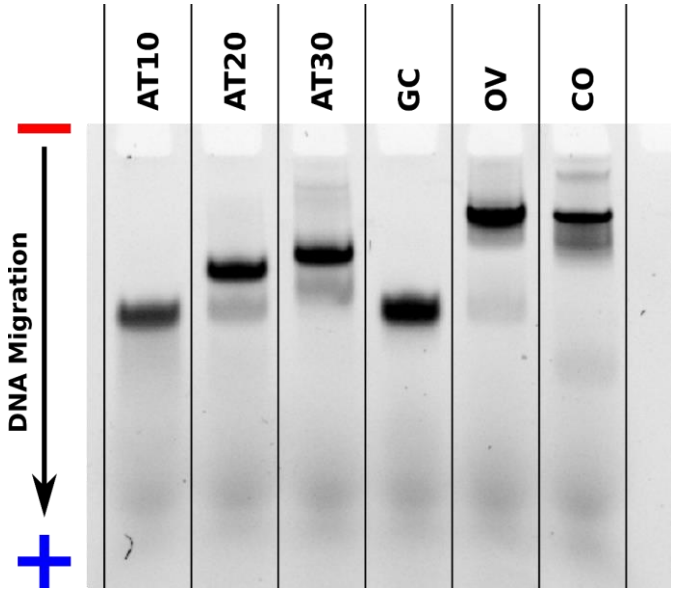

Figure S1. Gel electrophoresis of DX-tile folding. Folding of the DX-tile structures in a HEPES-Mg buffer (10 mM HEPES, 5 mM MgCl2, pH 7.0) analyzed *via* 10% polyacrylamide in 1X TBE gel electrophoresis (PAGE). PAGE gels were pre-run for 15 min with HEPES-Mg buffer. After pre-running 1 pmol of DX-tile was loaded and the gel was run on ice for 10 min at 70 volts after which the voltage was increased to 120 volts for 40 minutes. Upon completion, 1X GelRed® was used to visualize DNA band migration.

AT10

Name	Sequence (5' – 3')	Source	T_m (°C)	T_{ms} (°C)
AT10-1	TCAGAAGAACAAAAAAAAAA	IDT	39	35
AT10-2	ACGCATTTGTCGTTCTTCTGATGACCTTGCA	IDT	62	59
AT10-3	GTACACCTCGCTTTTTTTTTTGTGTCTGCAA	IDT	59	57
XAT-4	TTGCAGACACGACAAATGCGT	IDT	52	48
XAT-5	TGCAAGGTCCGCGAGGTGTAC	IDT	58	54
XAT-4-405	/5Alex405N/TTGCAGACACGACAAATGCGT	Biosyn	52	48
XAT-5-647	TGCAAGGTCCGCGAGGTG/iAlex647N/TAC	IDT	58	54

T_m = melting temperature, T_{ms} = sodium adjusted melting temperature

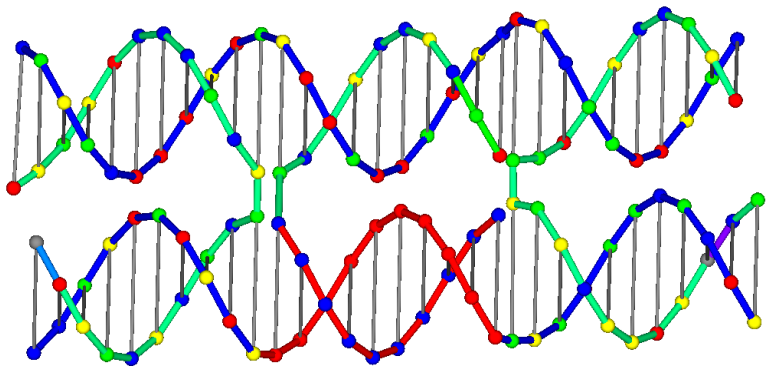

Figure S2. AT10 DNA sequences and structures schematic. Base key corresponds to adenine as blue, thymine as red, cytosine as yellow and guanine as green. Red regions indicates AT-track region. T_m and salt adjusted T_{ms} were calculated using the equations detailed above in the Supplementary Methods. These are theoretical melting temperatures of each strand hybridized to the complementary strand.

AT20

Name	Sequence (5′ – 3′)	Source	T_m (°C)	T_{ms} (°C)
AT20-1	GTACACCTCGCGTTTTTTTTTTTTTTTTTTTTGTGTCTGCAA	IDT	62	60
AT20-2	ACGCATTTGTCGTTCTTCTGACCTATGCGCCATGACCTTGCA	IDT	69	68
AT20-3	AGAAGAACAAAAAAAAAAAAAAAAAAAACTGGCGCATAGGTC	IDT	61	59
XAT-4	TTGCAGACACGACAAATGCGT	IDT	52	48
XAT-5	TGCAAGGTCCGCGAGGTGTAC	IDT	58	54
XAT-4-405	/5Alex405N/TTGCAGACACGACAAATGCGT	Biosyn	52	48
XAT-5-647	TGCAAGGTCCGCGAGGTG/iAlex647N/TAC	IDT	58	54

T_m = melting temperature, T_{ms} = sodium adjusted melting temperature

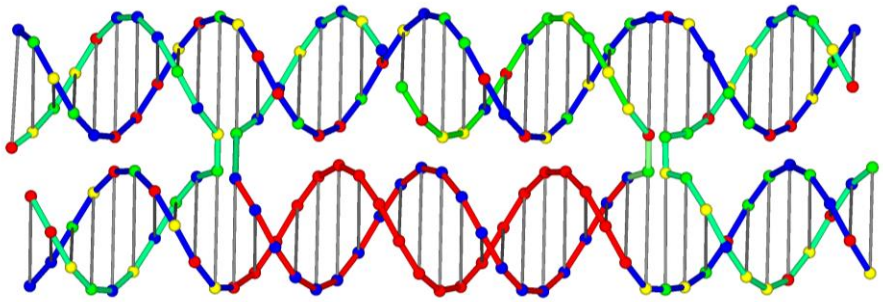

Figure S3. AT20 DNA sequences and structures schematic. Base key corresponds to adenine as blue, thymine as red, cytosine as yellow and guanine as green. Red regions indicates AT-track region. T_m and salt adjusted T_{ms} were calculated using the equations detailed above in the Supplementary Methods. These are theoretical melting temperatures of each strand hybridized to the complementary strand.

AT30

Name	Sequence (5'–3')	Source	T_m (°C)	T_{ms} (°C)
AT30-1	AGAAGAACAAAAAAAAAAAAAAAAAAAAAAAAAAACTCAGGGTCTCTGGCGCATAGGTC	IDT	66	66
AT30-2	GTACACCTCGCGTTTTTTTTTTTTTTTTTTTTTTTTTTTTGTGTCTGCAA	IDT	62	62
AT30-3	ACGCATTTGTCGTTCTTCTGACCTATGCGCCAGAGACCCTGATGACCTTGCA	IDT	73	73
XAT-4	TTGCAGACACGACAAATGCGT	IDT	52	48
XAT-5	TGCAAGGTCCGCGAGGTGTAC	IDT	58	54
XAT-4-405	/5Alex405N/TTGCAGACACGACAAATGCGT	Biosyn	52	48
XAT-5-647	TGCAAGGTCCGCGAGGTG/iAlex647N/TAC	IDT	58	54

T_m = melting temperature, T_{ms} = sodium adjusted melting temperature

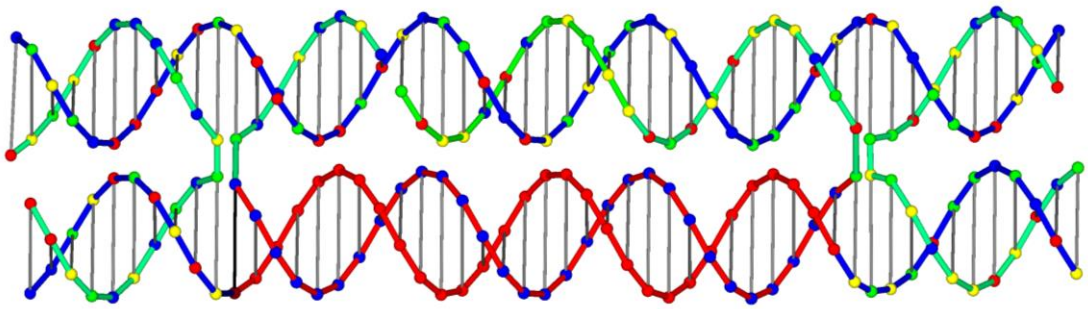

Figure S4. AT30 DNA sequences and structures schematic. Base key corresponds to adenine as blue, thymine as red, cytosine as yellow and guanine as green. Red regions indicates AT-track region. T_m and salt adjusted T_{ms} were calculated using the equations detailed above in the Supplementary Methods. These are theoretical melting temperatures of each strand hybridized to the complementary strand.

Overlapping (OV)

Name	Sequence (5'–3')	Source	T_m (°C)	T_{ms} (°C)
OV-1	GCAAAAAAAAAATCTATAAGCGCAGGGT	IDT	56	53
OV-2	AGGCTGCGGTGGTACACCTCACTTATAGATTTTTTTTTTGCGCTTTTTTTTTTGTGTCTGCAA	IDT	70	70
OV-3	ACGCATTTGTCGTTCTTCTTTTTTTTGCCACAGAGACCCTGCTGACCTTGCAATATCTCGCCAG	IDT	74	74
OV-4	CTCTGTGGCAAAAAAAAGAAGAACAAAAAAAAAAGC	IDT	59	57
XAT-4	TTGCAGACACGACAAATGCGT	IDT	52	48
OV-6	TGAGGTGTACCACCGCAGCCTCTGGCGAGATATTGCAAGGTC	IDT	72	71
XAT-4-405	/5Alex405N/TTGCAGACACGACAAATGCGT	Biosyn	52	48
OV-6-647	/5Alex647N/TGAGGTGTACCACCGCAGCCTCTGGCGAGATATTGCAAGGTC	IDT	72	71

T_m = melting temperature, T_{ms} = sodium adjusted melting temperature

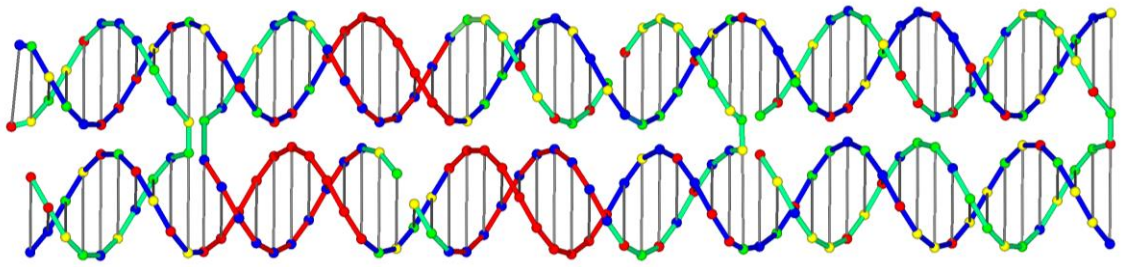

Figure S5. OV DNA sequences and structures schematic. Base key corresponds to adenine as blue, thymine as red, cytosine as yellow and guanine as green. Red regions indicates AT-track region. T_m and salt adjusted T_{ms} were calculated using the equations detailed above in the Supplementary Methods. These are theoretical melting temperatures of each strand hybridized to the complementary strand.

Continuous (CO)

Name	Sequence (5' – 3')	Source	T_m (°C)	T_{ms} (°C)
CO-1	CAAAAAAAAAACGCTCAGGGTCTCTGAAAAAAAAATC	IDT	60	58
CO-2	AGGCTGCGGTGGTACACCTCGCGTTTTTTTTTTGGCACT GCGCTTTTTTTTTTGTGTCTGCAA	IDT	73	73
CO-3	ACGCATTTGTCGTTCTTCTGATTTTTTTTTCAGAGACCCT GATGACCTTGCAATATCTCGCCAG	IDT	72	72
XAT-4	TTGCAGACACGACAAATGCGT	IDT	52	48
CO-5	AGAAGAACAAAAAAAAAAGCGCAGTGC	IDT	55	52
CO-6	TACCACCGCAGCCTCTGGCGAGATATTGCAAGGTCCGA GGTG	IDT	73	72
XAT-4-405	/5Alex405N/TTGCAGACACGACAAATGCGT	Biosyn	55	52
CO-6-647	/5Alex647N/TACCACCGCAGCCTCTGGCGAGATATTGC AAGGTCCGAGGTG	IDT	73	72

T_m = melting temperature, T_{ms} = sodium adjusted melting temperature

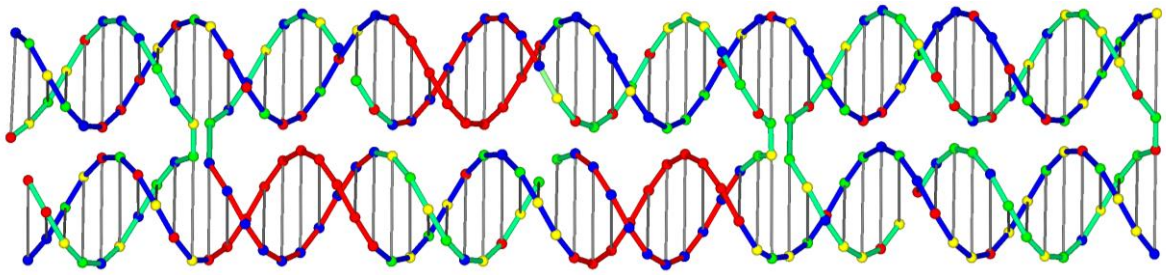

Figure S6. ATCO DNA sequences and structures schematic. Base key corresponds to adenine as blue, thymine as red, cytosine as yellow and guanine as green. Red regions indicates AT-track region. T_m and salt adjusted T_{ms} were calculated using the equations detailed above in the Supplementary Methods. These are theoretical melting temperatures of each strand hybridized to the complementary strand.

GC

Name	Sequence (5' – 3')	Source	T_m (°C)	T_{ms} (°C)
GC-1	TCAGAAGAACGCGCGCGCGC	IDT	60	55
AT10-2	ACGCATTTGTCGTTCTTCTGATGACCTTGCA	IDT	62	59
GC-3	GTACACCTCGAGCGCGCGCGCTTGTCTGCAA	IDT	70	67
GC-4	TTGCAGACAAGACAAATGCGT	IDT	50	46
XAT-5	TGCAAGGTCCGCGAGGTGTAC	IDT	58	54
GC-4-405	/5Alex405N/TTGCAGACACGACAAATGCGT	Biosyn	50	46
XAT-5-647	TGCAAGGTCCGCGAGGTG/iAlex647N/TAC	IDT	58	54

T_m = melting temperature, T_{ms} = sodium adjusted melting temperature

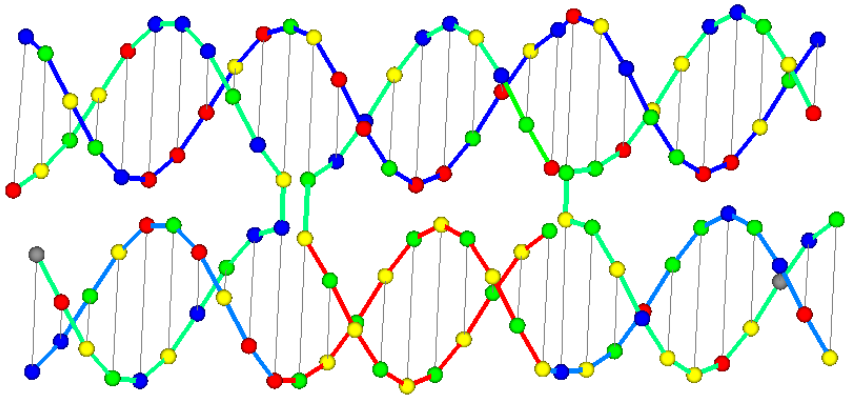

Figure S7. GC DNA sequences and structures schematic. Base key corresponds to adenine as blue, thymine as red, cytosine as yellow and guanine as green. Red regions indicates AT-track region. T_m and salt adjusted T_{ms} were calculated using the equations detailed above in the Supplementary Methods. These are theoretical melting temperatures of each strand hybridized to the complementary strand.

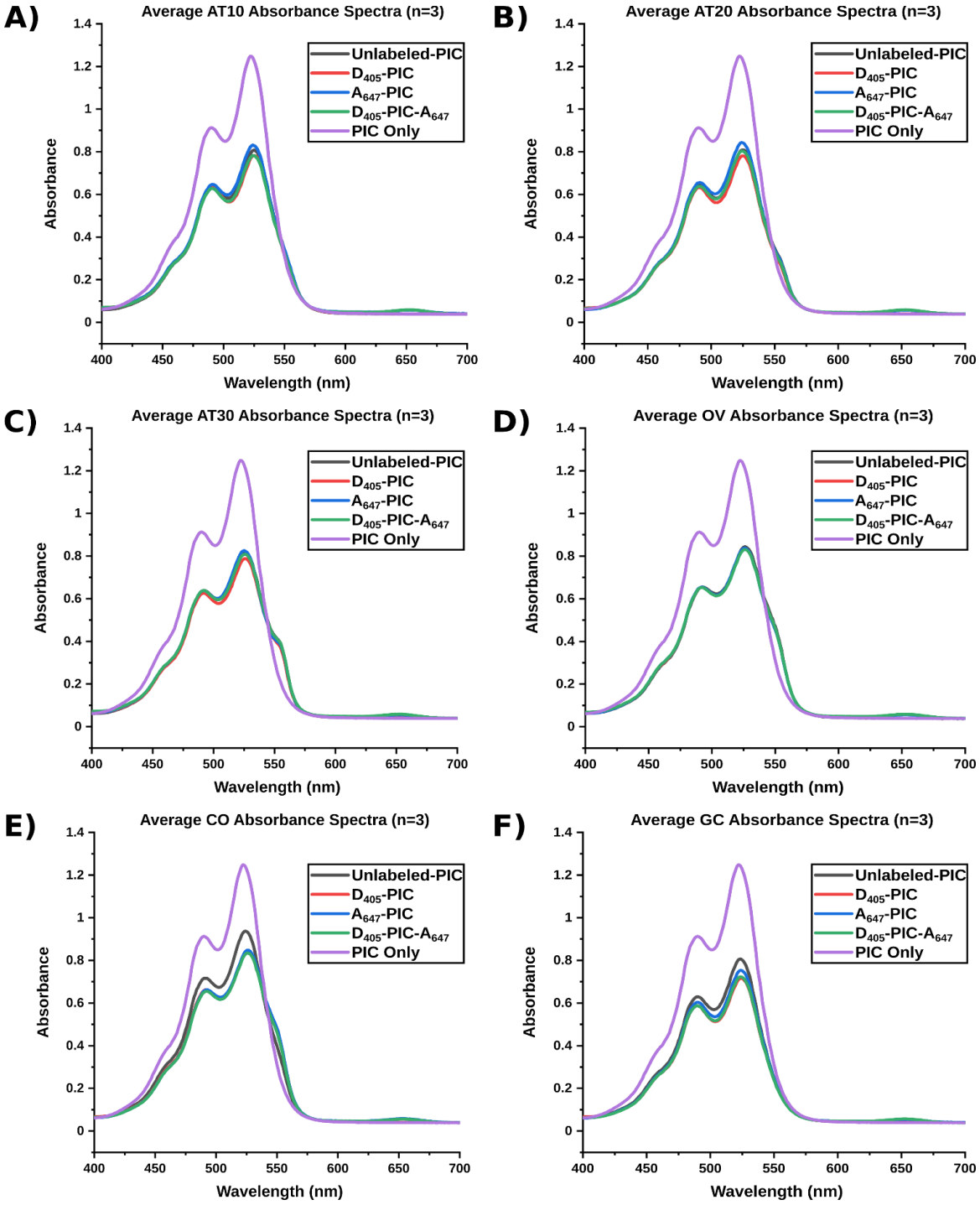

Figure S8. Averaged absorbance spectra of unlabeled-PIC, D_{405}-PIC, A_{647}-PIC, and D_{405}-PIC-A_{647} labeled samples compared to a PIC only negative control with no DNA. Contiguous AT-track (A-C), non-contiguous AT-track (D-E), and GC (F) scaffolds were all measured at 400 nM in the presence of 52 µM PIC. The decrease in monomer band absorbance has been attributed to characteristics of PIC dimerization.

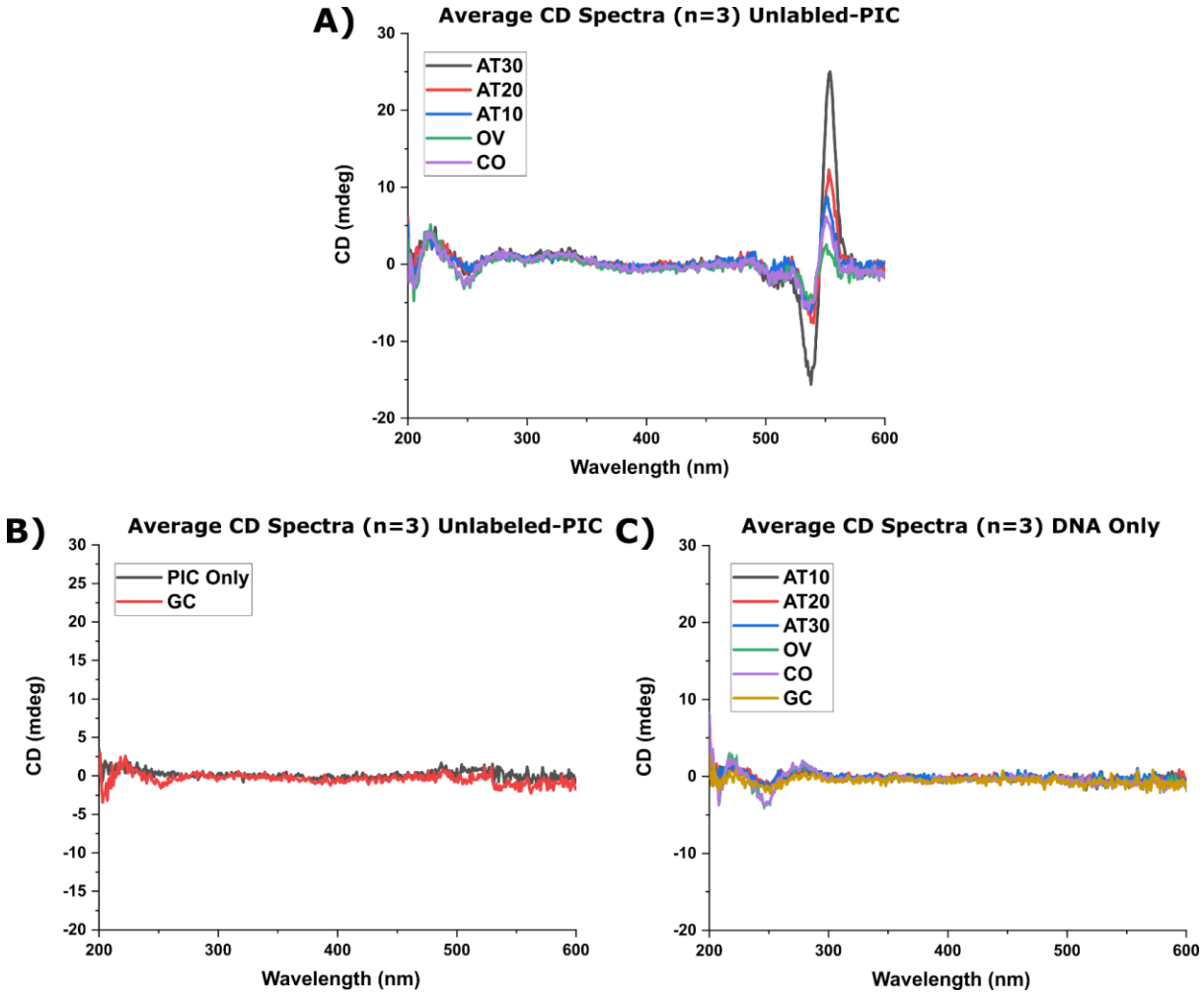

Figure S9. Average circular dichroism spectra of (A) AT-track containing scaffolds in the presence of PIC, (B) GC scaffold in the presence of PIC compared to a PIC only control with no DNA present, (C) all unlabeled DNA scaffolds in absence of PIC (*i.e.* DNA only). In all samples DNA is 400 nM and PIC is 52 µM.

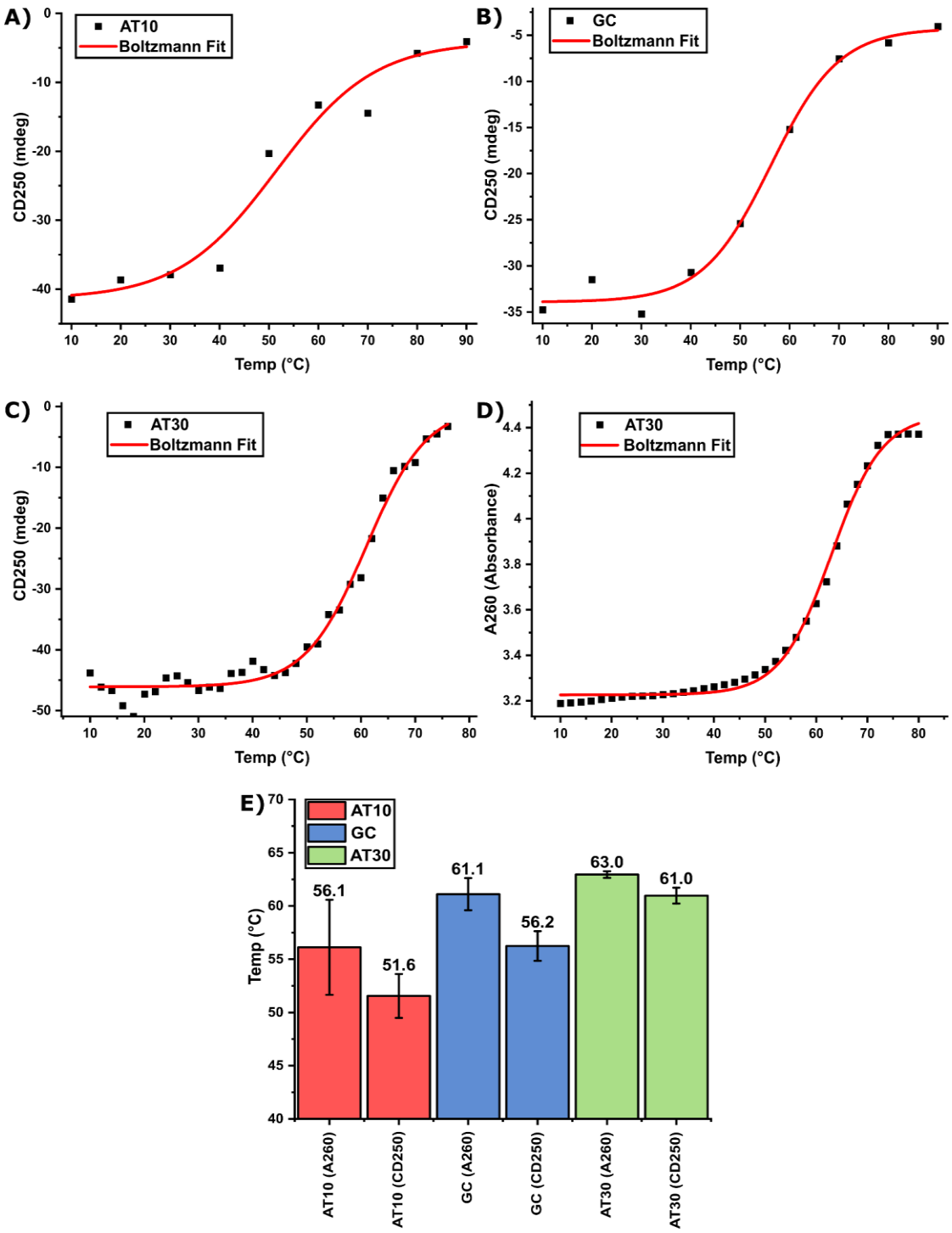

Figure S10. (A-C) CD measurements at 250 nm as a function of sample temperature for the **(A)** AT10, **(B)** GC, and **(C)** AT30 (black) as well as a Boltzmann fit (red). **(D)** Absorbance measurements at 260 nm for the AT30 structure (black) as well as a Boltzmann fit (red). **(E)** Comparison of the AT10, GC and AT30 melting temperatures using the Boltzmann function to fit CD250 and A260 melting curves. All melting curves were measured using 2.5 μM DX-tile.

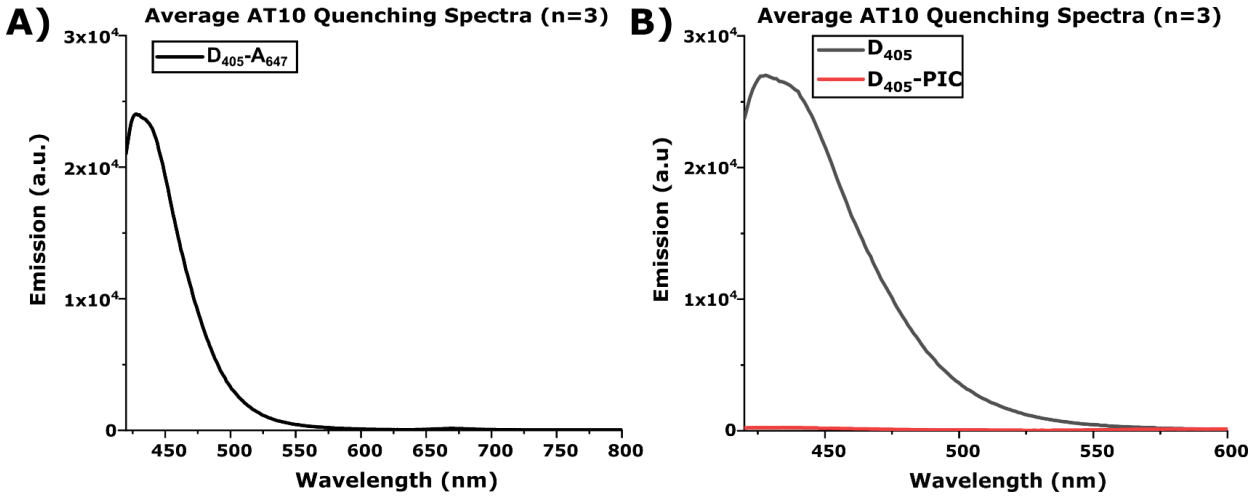

Figure S11. (A) Averaged AT10 emission spectra of D_{405} and A_{647} labeled scaffold in the absence of PIC to show no detectable direct transfer. AT10 has the smallest D405-A647 separation distance and hence most likely to engage in direct transfer. **(B)** Average AT10 emission spectra of D_{405} only and D_{405}-PIC labeled scaffolds to show D_{405} emission quenching by J-bit. Scaffolds were all measured at 400 nM and 52 µM PIC where applicable. All samples excited with 395 nm laser.

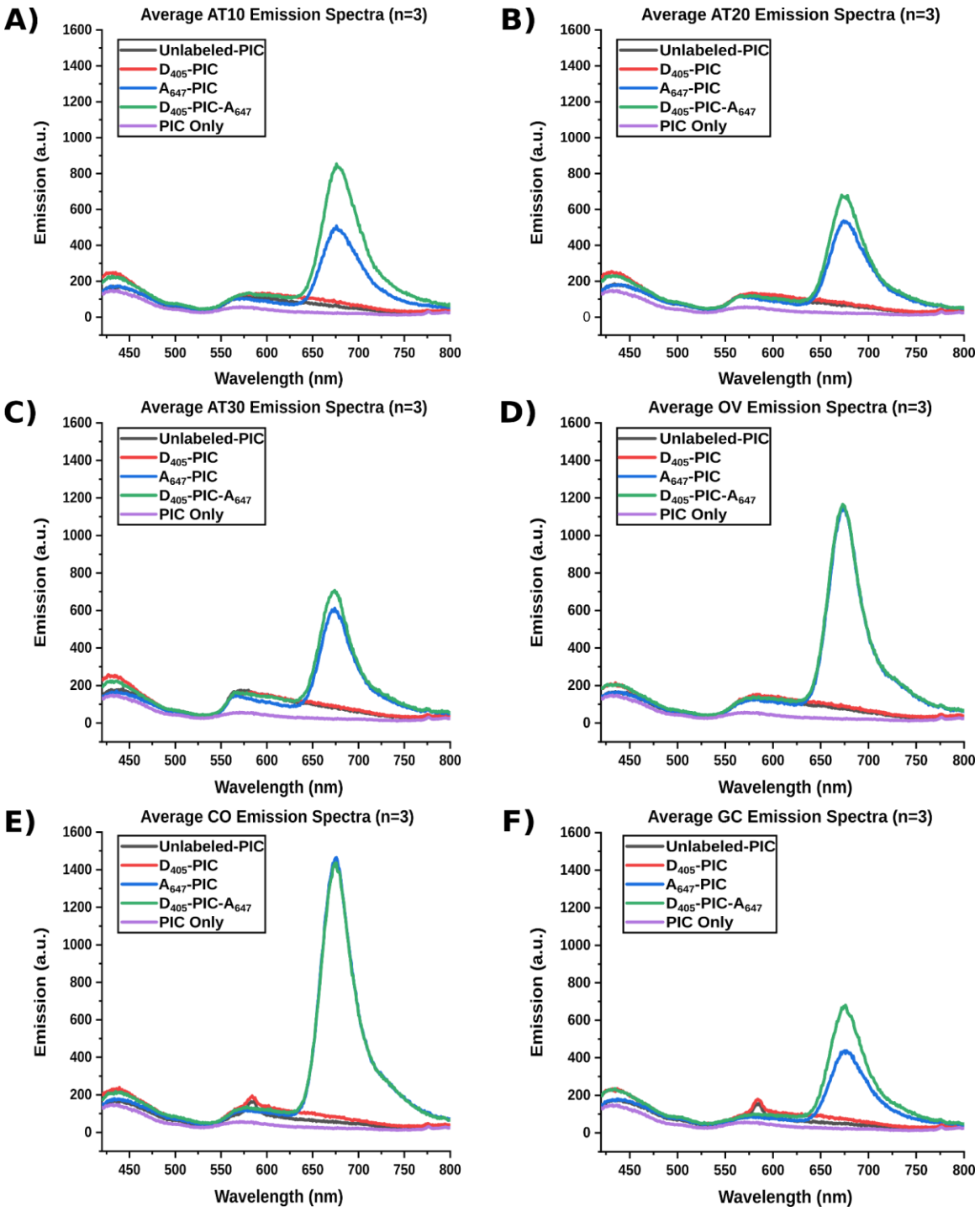

Figure S12. Averaged emission spectra of unlabeled-PIC, D_{405}-PIC, PIC-A_{647}, and D_{405}-PIC-A_{647} labeled samples compared to a PIC only negative control with no DNA. Contiguous AT-track (A-C), non-contiguous AT-track (D-E), and GC (F) scaffolds were all measured at 400 nM in the presence of 52 µM PIC using an excitation wavelength of 395 nm.

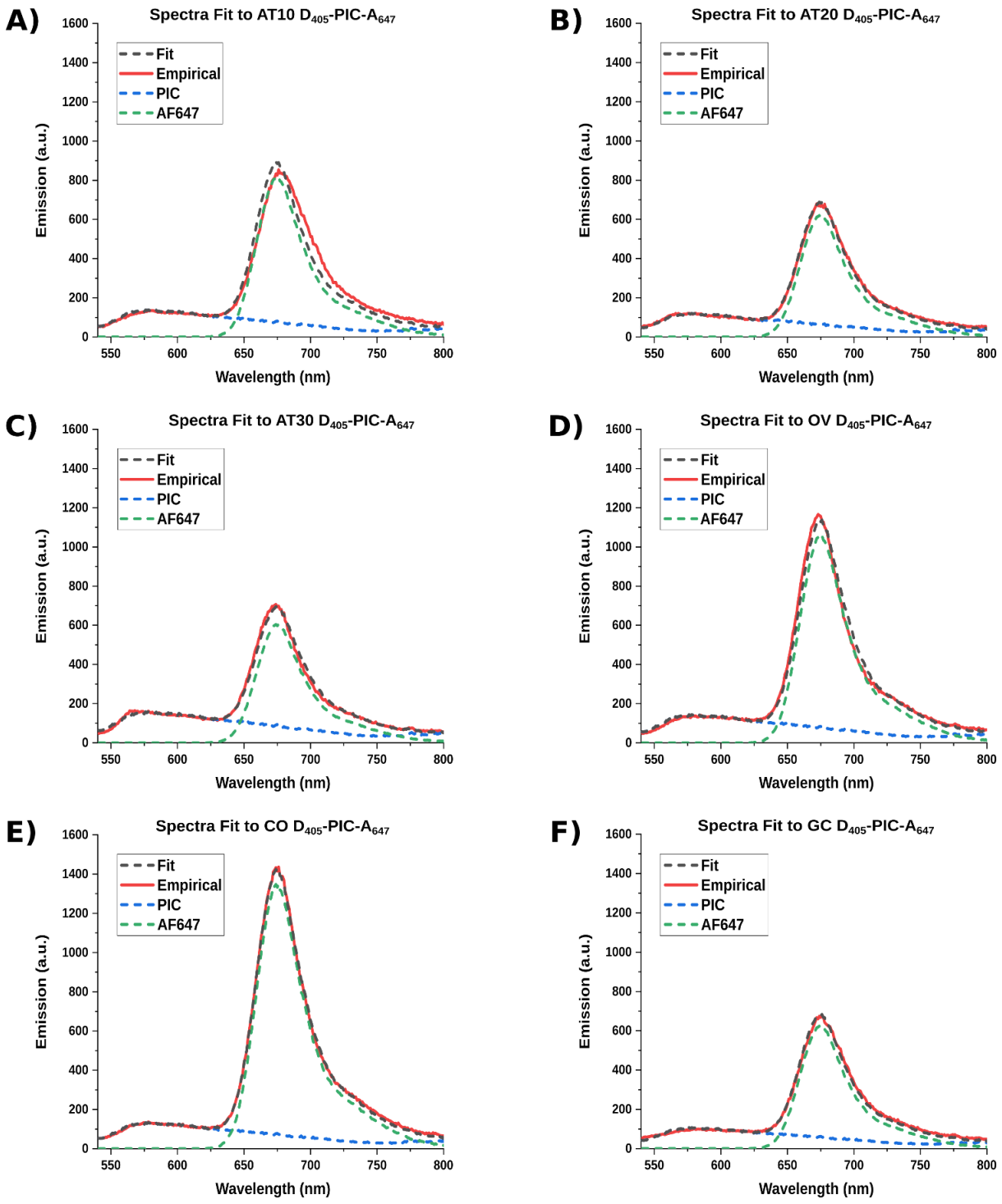

Figure S13. Spectral decomposition plots of the D_{405}-PIC-A_{647} emission curve for each scaffold. Contiguous AT-track (A-C), non-contiguous AT-track (D-E), and GC (F) scaffold emission curves were fit using the A|E UV-Vis-IR Spectral Software as outlined in the supporting methods section above. The fit spectrum is a linear combination of the scaled components.

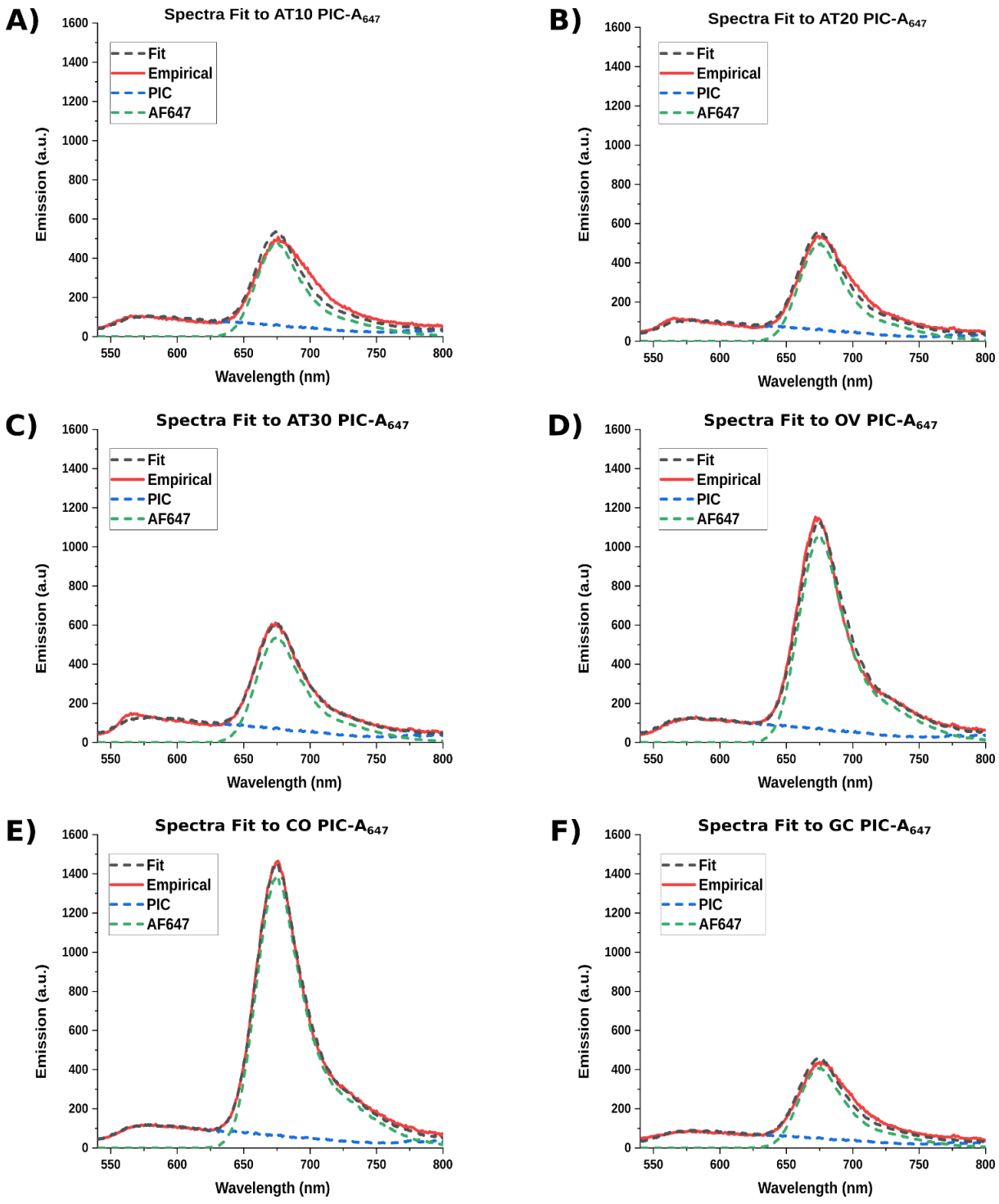

Figure S14. Spectral decomposition plots of the PIC-A_{647} emission curve for each scaffold. Contiguous AT-track (A-C), non-contiguous AT-track (D-E), and GC (F) scaffold emission curves were fit using the A|E UV-Vis-IR Spectral Software as outlined in the supporting methods section above. The fit spectrum is a linear combination of the scaled components.

Table S1. Goodness of fit values for Spectral decompositions.

	AT10	AT20	AT30	OV	CO	GC
AF647 Only	0.959	0.982	0.995	0.994	0.999	0.978
Full	0.967	0.998	0.993	0.992	0.999	0.996

R^2 values were calculated by taking the ratio of the residual sum of squares over the total sum of squares.

Supporting References

1. Förster, T., Zwischenmolekulare Energiewanderung und Fluoreszenz. *Annalen Der Physik* **1948,** *437* (1-2), 55-75.

2. Lakowicz, J. R., *Principles of Fluorescence Spectroscopy*. Springer Science & Business Media: **2013**.

3. Algar, W. R.; Hildebrandt, N.; Vogel, S. S.; Medintz, I. L., FRET as a Biomolecular Research Tool—Understanding Its Potential While Avoiding Pitfalls. *Nat. Methods.* **2019,** *16* (9), 815-829.

4. Medintz, I. L.; Hildebrandt, N., *FRET-Förster Resonance Energy Transfer: From Theory to Applications*. John Wiley & Sons: **2013**.

5. Uv-Vis-Ir, A., Spectral Software 1.2, Fluortools.

6. Würthner, F.; Kaiser, T. E.; Saha-Möller, C. R., J-Aggregates: From Serendipitous Discovery to Supramolecular Engineering of Functional Dye Materials. *Angew. Chem. Int. Ed.* **2011,** *50* (15), 3376-3410.

7. Ware, W. R.; Rothman, W., Relative Fluorescence Quantum Yields Using an Integrating Sphere. The Quantum Yield of 9, 10-Diphenylanthracene in Cyclohexane. *Chem. Phys. Lett.* **1976,** *39* (3), 449-453.

8. Sens, R.; Drexhage, K. H., Fluorescence Quantum Yield of Oxazine and Carbazine Laser Dyes. *J. Lumin.* **1981,** *24*, 709-712.

9. Kibbe, W. A., OligoCalc: An Online Oligonucleotide Properties Calculator. *Nucleic Acids Res.* **2007**, 35 (suppl_2), W43-W46.