SUPPORTING INFORMATION

Machine learning assembly landscapes from particle tracking data

Andrew W. Long[1], Jie Zhang[1], Steve Granick[2], and Andrew L. Ferguson[1,*]

[1]Department of Materials Science and Engineering, University of Illinois at Urbana-Champaign, Urbana IL 61801, USA

[2]Center for Soft and Living Matter, Ulsan National Institute of Science and Technology, 50 UNIST-gil, Ulju-gun, Ulsan 689-798, South Korea

[*]Corresponding author. Email: alf@illinois.edu. Tel: (217) 300-2354. Fax: (217) 333-2736.

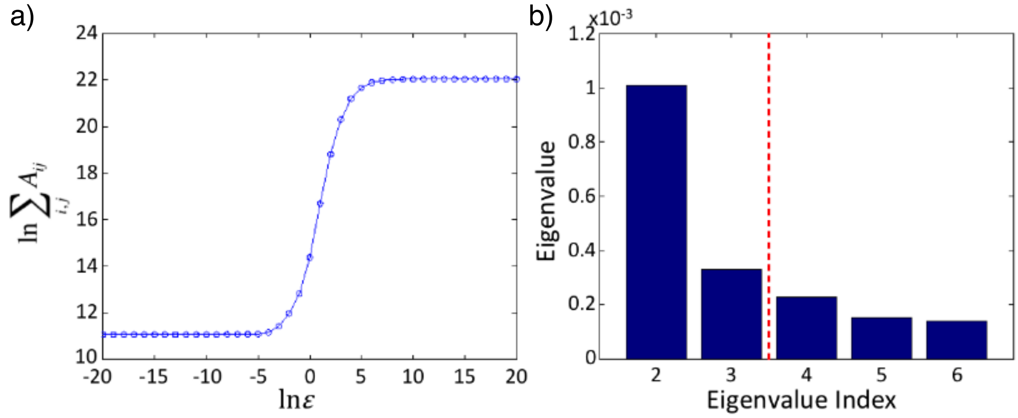


Figure S1: Specification of diffusion map bandwidth and effective dimensionality for the self-assembly of the mixture of active Janus particles and passive linker particles. (a) Specification of the diffusion map Gaussian kernel bandwidth using the approach in Ref. [1]. A log-log plot of the sum of the elements of the pairwise similarity matrix, $\sum_{i,j} A_{ij}$, as a function of Gaussian kernel bandwidth, ϵ, exhibits a sigmoid shape. The lower plateau defines ϵ values for which the discrete Markov chain is effectively disconnected prohibiting the construction of a unified diffusion map embedding. The upper plateau defines ϵ values for which the Markov chain is fully connected, precluding discrimination between local and non-local hops, and therefore fast and slow time scales. Accordingly, an appropriate bandwidth should be selected from the linear region. In practice, we find better results from selections near the intersection of the linear region and the upper plateau, motivating our choice of $\epsilon = \exp(10)$. (b) Diffusion map eigenvalue spectrum omitting the trivial unit eigenvalue, $\lambda_1 = 1$, for viewing clarity. Application of the L-method of Salvador and Chan[2] identifies the first “knee” (i.e., discontinuity) in the eigenvalue spectrum between λ_3 and λ_4, implying a spectral gap after the second non-trivial eigenvalue and an effective dimensionality of $k = 2$. This motivated us to construct diffusion map embeddings into the top two non-trivial eigenvectors (Ψ_2, Ψ_3).

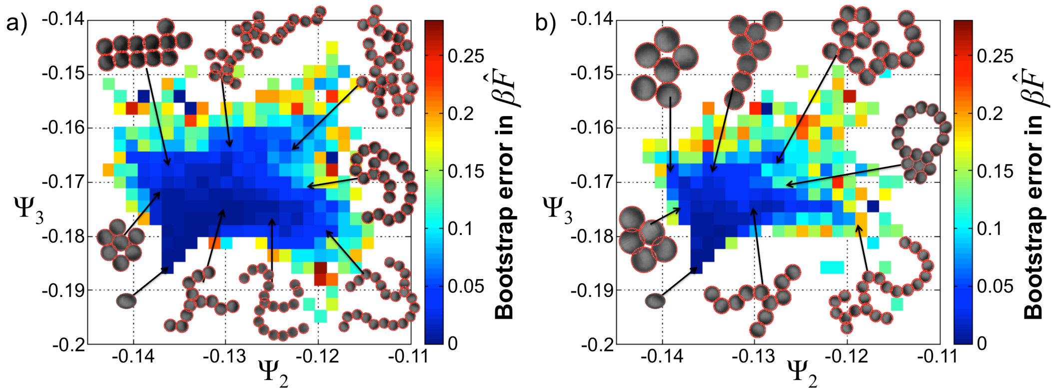


Figure S2: Estimated uncertainties in effective free energy surfaces for the self-assembly of a mixture of active Janus particles with passive linker particles presented in Figure 3c,d estimated by 10 bootstrap resamples at high field strength E (583-833 V/cm) and (a) low AC frequency f (70-300 kHz) and (b) high AC frequency (600 kHz-5MHz).

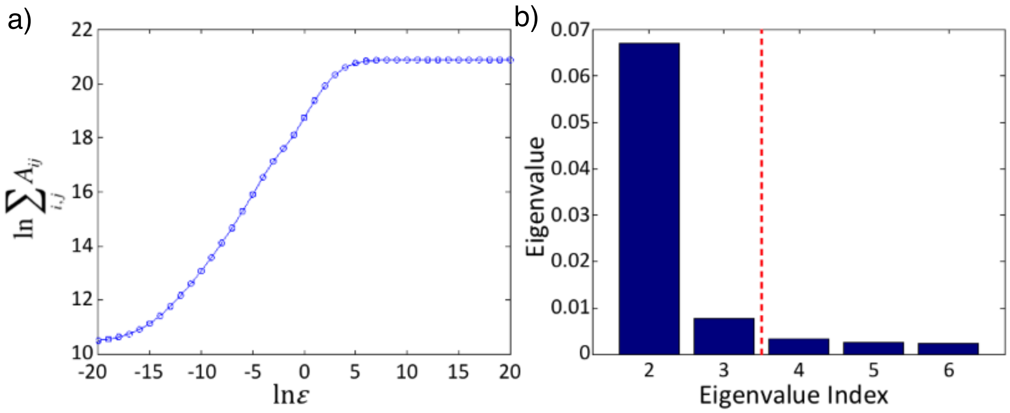


Figure S3: Specification of diffusion map bandwidth and effective dimensionality for the self-assembly of active Janus particles. (a) Specification of the diffusion map Gaussian kernel bandwidth using the approach in Ref. [1]. We selected a value of $\epsilon = \exp(5.5)$. (b) Diffusion map eigenvalue spectrum omitting the trivial unit eigenvalue, $\lambda_1 = 1$, for viewing clarity. Application of the L-method of Salvador and Chan[2] identifies the first "knee" (i.e., discontinuity) in the eigenvalue spectrum between λ_3 and λ_4, implying a spectral gap after the second non-trivial eigenvalue and an effective dimensionality of $k = 2$. This motivated us to construct diffusion map embeddings into the top two non-trivial eigenvectors (Ψ_2, Ψ_3).

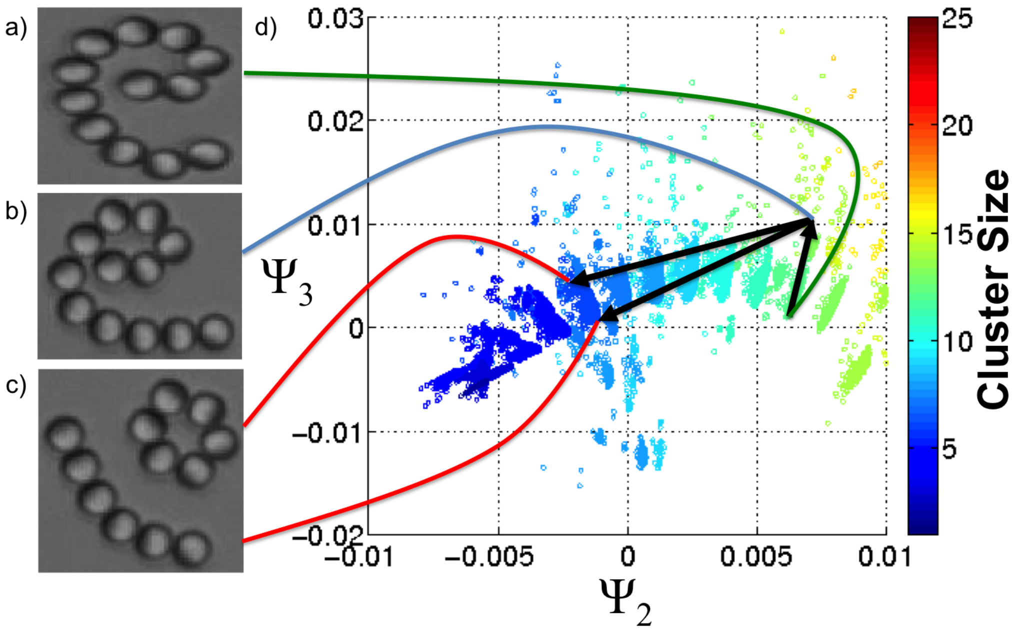


Figure S4: An example of a ring formation event extracted from the particle tracking trajectory for the self-assembly of active Janus particles at E=583.3 V/cm, f=1.5 MHz, [NaCl] = 0.01mM traced over the composite intrinsic manifold discovered by the diffusion map (Figure 5). (a) A linear 12-mer chain progressing under rICEP coils back upon itself. (b) The head of the chain collides with the middle of its own tail leading to the formation of a 6-mer ring and a residual 6-mer linear chain. (c) The ring and chain disaggregate leading to the formation of two new and distinct clusters. (d) The progression of the ring formation event superposed onto the composite intrinsic manifold. The coiling back of the 12-mer chain results in a local step over the manifold, followed by a non-local jump corresponding to the splitting of the 12-mer into two 6-mers. Internal cluster rearrangements and monomer addition or removal corresponds to local moves over the manifold, whereas large scale cluster aggregation or disaggregation events correspond to non-local jumps.

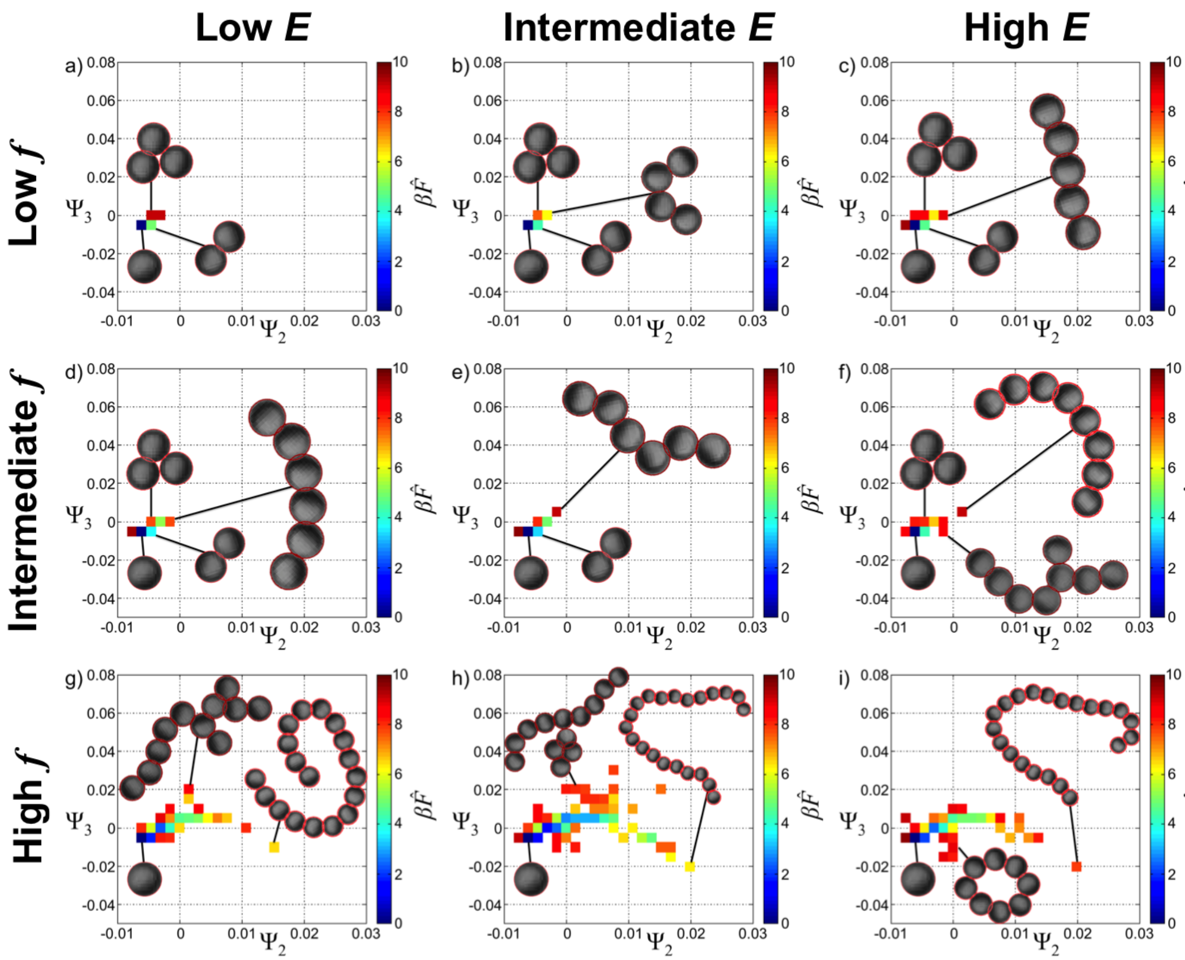


Figure S5: Effective free energy landscapes at high (0.1 mM) NaCl concentration at different applied AC electric field strengths, E, and frequencies, f. Columns partition the electric field strength into low (167-333 V/cm), intermediate (417-583 V/cm), and high (667-833 V/cm) regimes. Rows split the AC frequency of the applied field into low (50-300 kHz), intermediate (400-800 kHz), and high (900 kHz - 3 MHz) regimes. Assembly is arrested under low and intermediate frequencies under these salt conditions, with high frequency behavior analogous to the low salt system.

	Field Strength (V/cm)					
Frequency (kHz)	250	417	500	583	750	833
70	0	0	0	0	0	1
100	1	3	1	1	1	5
150	0	0	0	0	0	1
200	0	0	0	0	0	3
300	0	1	0	0	0	1
600	0	0	0	1	0	2
1000	0	0	1	0	0	1
5000	0	0	0	0	0	1
7000	0	0	0	0	0	1
9000	0	0	0	0	0	1
11000	0	0	0	0	0	1

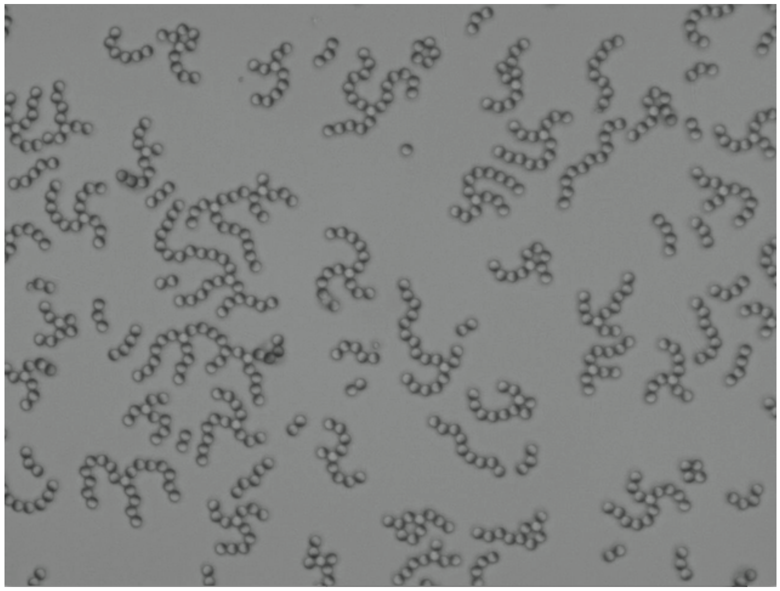
Table S1: Number of experiments performed for a self-assembling system of active metallodielectric Janus particles with passive linker particles at each combination of field strength and AC frequency at a salt concentration of 0.1mM NaCl, described in Section 3.1.

0.01mM	Electric Field Strength E (V/cm)								
Frequency f (kHz)	167	250	333	417	500	583	667	750	833
50	0	1	1	1	1	1	1	1	1
100	1	1	1	1	1	1	1	1	1
120	0	0	0	0	0	0	0	0	0
150	0	0	0	0	0	0	0	0	0
200	1	1	1	1	1	1	1	1	1
300	0	1	1	1	1	1	1	1	1
400	0	0	0	1	0	0	0	0	1
500	1	1	1	1	1	1	1	1	1
600	1	1	1	1	1	1	1	1	1
700	1	1	1	1	1	1	1	1	1
800	1	1	1	1	1	1	1	1	1
900	1	1	1	1	1	1	1	1	1
1000	1	1	1	1	1	1	1	1	1
1500	1	1	1	1	1	1	1	1	1
2000	1	1	1	1	1	1	1	1	1
2500	1	1	1	1	1	1	1	1	1
3000	1	1	1	1	1	1	1	1	1
4000	0	1	1	1	1	1	1	1	1
5000	0	0	0	1	0	0	0	0	1
6000	1	1	1	1	1	1	1	1	1
7000	0	0	0	1	0	0	0	0	1
8000	1	1	1	1	1	1	1	1	1
9000	0	0	0	0	0	0	0	0	1
10000	0	0	0	0	0	0	0	0	0

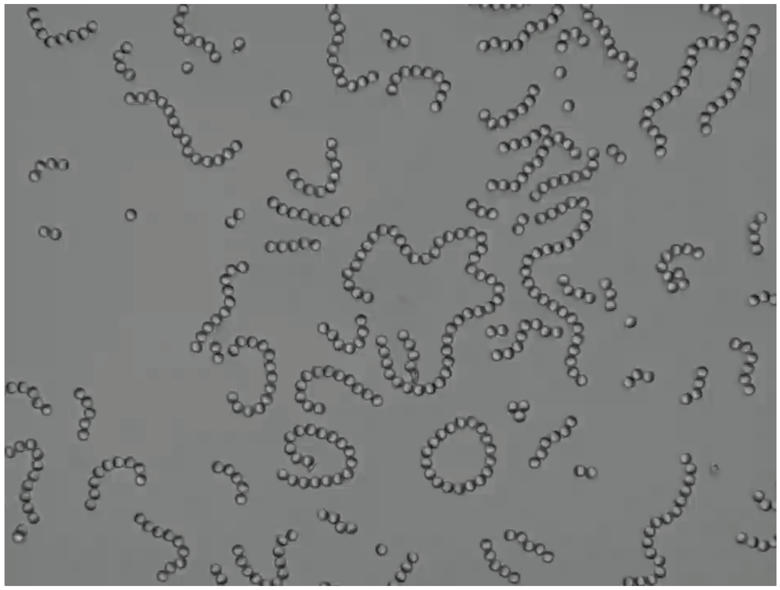
Table S2: Number of experiments performed for a homogeneous self-assembling system of active metallodielectric Janus particles at a salt condition of 0.01mM NaCl, described in Section 3.2.

0.1mM	Electric Field Strength E (V/cm)								
Frequency f (kHz)	167	250	333	417	500	583	667	750	833
50	0	0	0	0	0	0	0	0	0
100	1	1	1	1	1	1	1	1	1
120	0	1	1	1	1	1	1	1	1
150	1	1	1	1	1	1	1	1	1
200	1	1	1	1	1	1	1	1	1
300	1	1	1	1	1	1	1	1	1
400	1	1	1	1	1	1	1	1	1
500	1	1	1	1	1	1	1	1	1
600	1	1	1	1	1	1	1	1	1
700	1	1	1	1	1	1	1	1	1
800	1	1	1	1	1	1	1	1	1
900	1	1	1	1	1	1	1	1	1
1000	1	1	1	1	1	1	1	1	1
1500	1	1	1	1	1	1	1	1	1
2000	1	1	1	1	1	1	1	1	1
2500	0	0	0	0	0	0	0	0	0
3000	1	1	1	1	1	1	1	1	1
4000	1	1	1	1	1	1	1	1	1
5000	1	1	1	1	1	1	1	1	1
6000	1	1	1	1	1	1	1	1	1
7000	1	1	1	1	1	1	1	1	1
8000	1	1	1	1	1	1	1	1	1
9000	1	1	1	1	1	1	1	1	1
10000	0	0	0	0	0	0	0	0	1

Table S3: Number of experiments performed for a homogeneous self-assembling system of active metallodielectric Janus particles at a salt condition of 0.1mM NaCl, described in Section 3.2.



Movie S1: Experimental self-assembly trajectory for a mixture of active metallodielectric Janus particles with passive linker particles, under applied AC frequency $f = 100$ kHz and field strength $E = 833$ V/cm.



Movie S2: Experimental self-assembly trajectory for a homogenous system of active metallodielectric Janus particles under applied AC frequency $f = 1.5$ MHz and field strength $E = 583$ V/cm.

References

[1] Coifman, R.; Shkolnisky, Y.; Sigworth, F.; Singer, A. *IEEE Trans. Image Process.* **2008**, *17*, 1891–1899.

[2] Salvador, S.; Chan, P. Determining the number of clusters/segments in hierarchical clustering/segmentation algorithms. Tools with Artificial Intelligence, 2004. ICTAI 2004. 16th IEEE International Conference on. 2004; pp 576–584.