Supplemental Materials for "Self-Assembly of a Dimer System"

Contents

1 Link to Supplementary Code

IPython code for creating Fig. 4, Fig. 5, and for the biophysics calculations in Sec. VI in the main text can be found at
https://github.com/mowillia/dimer_self_assembly_code.

2 Deriving $a_{n,\ell}$ as a Series and an Integral

We are seeking a formula that answers the following question:

> Given $2n$ distinguishable objects that are all initially paired in some way, what is the number of ways to form ℓ pairs such that none of these new pairs coincide with any original pairings?

We call this number $a_{n,\ell}$, and it is easy to see what its value should be for $\ell = n$ and $\ell = 1$. If we were to take $\ell = n$, we would have the case of the "bridge couples problem" and we should obtain the formula derived in [1]. If were were to take $\ell = 1$, we could infer that $a_{n,1} = 2n(2n-2)/2$ since there are $2n$ ways to select the first element, $2n-2$ ways to select an element that was not initially paired with this first element, and a factor of $1/2$ for double counting.

To find the general formula for $a_{n,\ell}$, we employ the inclusion-exclusion principle [2].

First, we establish some definitions. We define $|A_i|_{n,\ell}$ as the number of way to reform ℓ pairs, out of $2n$ initially paired elements, such that in the new set of pairs, we include the ith pair of the initial pairings. We in turn say that the quantity

$$|A_{i_1} \cap \cdots \cap A_{i_k}|_{n,\ell}, \tag{S1}$$

equals the size of the set where, out of $2n$ initially paired elements, we have formed $\ell \leq n$ new pairs which include the pairs $i_1, \ldots, i_k$ (for $k \leq \ell$) of the original pairings. By this definition, our desired quantity $a_{n,\ell}$ can be written as

$$\sum_{1 \leq i_1 < \cdots < i_\ell \leq n}^{n} |A^c_{i_1} \cap \cdots \cap A^c_{i_\ell}|_{n,\ell}, \tag{S2}$$

where A^c_k is the complement of A_k. Eq.(S2) is the total number of ways to reform ℓ pairs out of $2n$ initially paired elements such that none of the ℓ pairs is found in the initial pairings. Given that the intersection of complements is equal to the complement of the union, we have.

$$\begin{aligned}\sum_{1\le i_1<\cdots<i_\ell\le n}^{n} |A_{i_1}^c \cap \cdots \cap A_{i_\ell}^c|_{n,\ell} &= \sum_{1\le i_1<\cdots<i_\ell\le n}^{n} |\,(A_{i_1}\cup\cdots\cup A_{i_\ell})^c\,|_{n,\ell} \\ &= |\mathcal{S}|_{n,\ell} - \sum_{1\le i_1<\cdots<i_\ell\le n}^{n} |A_{i_1}\cup\cdots\cup A_{i_\ell}|_{n},\end{aligned} \tag{S3}$$

where we we defined $|\mathcal{S}|_{n,\ell}$ as the number of ways to create $\ell \le n$ pairs out of a set of $2n$ elements. Combinatorics tells us that $|\mathcal{S}|_{n,\ell}$ is

$$|\mathcal{S}|_{n,\ell} = \binom{2n}{2\ell}\frac{(2\ell)!}{2^\ell \ell!} = \binom{2n}{2\ell}(2\ell-1)!!, \tag{S4}$$

Now, to compute Eq.(S2), we must calculate the last quantity in Eq.(S3), and we do so by the inclusion-exclusion principle. By the principle, we have

$$\begin{aligned}&\sum_{1\le i_1<\cdots<i_\ell\le n}^{n} |A_{i_1}\cup\cdots\cup A_{i_\ell}|_{n,\ell} \\ &\quad= \sum_{i=1}^{n}|A_i|_{n,\ell} - \sum_{1\le i<j\le n}^{n}|A_i\cap A_j|_{n,\ell} \quad + \cdots + \sum_{1\le i_1<\cdots<i_\ell\le n}^{n}(-1)^{\ell-1}|A_{i_1}\cap\cdots\cap A_{i_\ell}|_{n,\ell}.\end{aligned} \tag{S5}$$

We recall that $|A_i|_{n,\ell}$ equals the number of way to reform ℓ pairs, out of $2n$ initially paired elements, such that in the new set of pairs, we include the ith pair of the initial pairings. Since the ith pair is fixed in this pairing, the number of ways to achieve this new pairing is simply the number of ways to form $\ell-1$ pairs out of a set of $2n-2$ elements. Thus we have

$$|A_i|_{n,\ell} = \binom{2n-2}{2\ell-2}(2\ell-2-1)!!. \tag{S6}$$

This quantity is independent of which i we choose, so, in Eq.(S5), the summation can be replaced with the factor $\binom{n}{1}$. Similarly, the quantity $|A_i\cap A_j|_{n,\ell}$ is the number of ways to choose ℓ pairs, out of $2n$ initially paired elements, such that we include the ith and jth pairs of the original pairing. Thus, we have

$$|A_i\cap A_j|_{n,\ell} = \binom{2n-4}{2\ell-4}(2\ell-4-1)!!, \tag{S7}$$

and the summation is replaced with the factor $\binom{n}{2}$. Following this pattern, we find that Eq.(S5) becomes

$$\sum_{1\le i_1<\cdots<i_\ell\le n}^{n} |A_{i_1}\cup\cdots\cup A_{i_\ell}|_{n,\ell} = \sum_{j=1}^{\ell}(-1)^{j-1}\binom{n}{j}\binom{2n-2j}{2\ell-2j}(2\ell-2j-1)!!. \tag{S8}$$

Finally, using Eq.(S4) in Eq.(S3), and noting that final result is our desired $a_{n,\ell}$, we have

$$a_{n,\ell} = \sum_{j=0}^{\ell}(-1)^j\binom{n}{j}\binom{2n-2j}{2\ell-2j}(2\ell-2j-1)!!. \tag{S9}$$

We can also write Eq.(S9) as an integral which will later allow us to write the partition function as a double integral. The first step is to rewrite the second combinatorial factor as

$$\binom{2n-2j}{2\ell-2j} = \frac{2^{n-j}(n-j)!}{(2n-2\ell)!(2\ell-2j)!}(2n-2j-1)!!. \tag{S10}$$

We then find

$$a_{n,\ell} = \sum_{j=0}^{\ell}(-1)^j \binom{n}{j} \frac{2^{n-j}(n-j)!}{(2n-2\ell)!(2\ell-2j)!}\,(2n-2j-1)!!\,(2\ell-2j-1)!!$$

$$= \frac{2^{n-\ell}}{(2n-2\ell)!}\frac{n!}{\ell!}\sum_{j=0}^{\ell}(-1)^j \frac{\ell!}{j!(\ell-j)!}(2n-2j-1)!!$$

$$= \frac{2^{n-\ell}(n-\ell)!}{(2n-2\ell)!}\frac{n!}{\ell!(n-\ell)!}\sum_{j=0}^{\ell}(-1)^j \binom{\ell}{j}\frac{2^{n-j}}{\sqrt{\pi}}\Gamma(n-j+1/2), \tag{S11}$$

Using the integral definition of the Gamma function, we obtain.

$$a_{n,\ell} = \frac{1}{(2n-2\ell-1)!!}\binom{n}{\ell}\frac{1}{\sqrt{\pi}}\int_0^\infty dt\, e^{-t}t^{-1/2}(2t)^n(1-1/2t)^\ell. \tag{S12}$$

3 Derivation of Non-Gendered Partition Function

In deriving the final form of the partition function for the non-gendered system, we begin with the partition function expressed as a summation over the total number of dimers and the total number of correct dimers:

$$Z_N(V,T,E_0,\Delta) = \sum_{k=0}^{N}\sum_{m=0}^{k}\binom{N}{m}a_{N-m,k-m}\,e^{\beta(kE_0+m\Delta)}\left(\frac{V}{\lambda_0^3}\right)^{2N-2k}\left(\frac{V}{(\lambda_0/\sqrt{2})^3}\right)^k \tag{S13}$$

Using the integral expression Eq.(S12), we find Eq.(S13) becomes

$$\begin{aligned} Z_N(V,T,E_0,\Delta) &= \frac{(V/\lambda_0^3)^{2N}}{\sqrt{\pi}}\int_0^\infty dx\,\frac{e^{-x}}{\sqrt{x}}\sum_{k=0}^{N}\sum_{m=0}^{k}\delta^k\eta^m\frac{1}{(2N-2k-1)!!}\binom{N}{m}\binom{N-m}{k-m}(2x)^{N-k}(2x-1)^{k-m} \\ &= \frac{(V/\lambda_0^3)^{2N}}{\sqrt{\pi}}\int_0^\infty dx\,\frac{e^{-x}}{\sqrt{x}}(2x)^N\sum_{k=0}^{N}\sum_{m=0}^{k}\frac{[\delta(1-1/2x)]^k}{(2N-2k-1)!!}\binom{N}{N-m}\binom{N-m}{k-m}\left(\frac{\eta}{2x-1}\right)^m \end{aligned} \tag{S14}$$

where we denoted

$$\delta \equiv \frac{2\sqrt{2}\lambda_0^3}{V}e^{\beta E_0}, \qquad \eta \equiv e^{\beta\Delta}. \tag{S15}$$

Next, we isolate the sum over m to find

$$\sum_{m=0}^{k}\binom{N}{N-m}\binom{N-m}{k-m}\left(\frac{\eta}{2x-1}\right)^m = \binom{N}{k}\left(\frac{\eta}{2x-1}+1\right)^k, \tag{S16}$$

where we used the fact that $\binom{n}{k}=0$ if $k<0$, and the identity $\sum_{k=0}^{n}\binom{n}{k}\binom{k}{r}x^k = x^r(1+x)^{n-r}\binom{n}{r}$. Returning to Eq.(S14), we find

$$\begin{aligned} Z_N(V,T,E_0,\Delta) &= \frac{(V/\lambda_0^3)^{2N}}{\sqrt{\pi}}\int_0^\infty dx\,\frac{e^{-x}}{\sqrt{x}}(2x)^N\sum_{k=0}^{N}\binom{N}{k}\frac{1}{(2N-2k-1)!!}\left[\frac{\delta}{2x}(\eta+2x-1)\right]^k \\ &= \frac{(V/\lambda_0^3)^N 2^N}{\sqrt{\pi}(2N-1)!!}\int_0^\infty dx\,\frac{e^{-x}}{\sqrt{x}}x^N\sum_{k=0}^{N}\binom{2N}{2k}(2k-1)!!\left[\frac{\delta}{2x}(\eta+2x-1)\right]^k. \end{aligned} \tag{S17}$$

Then, using the integral identity

$$\sum_{k=0}^{N}\binom{2N}{2k}(2k-1)!!\,\Lambda^k = \frac{1}{2\sqrt{\pi}}\int_0^\infty dy\,\frac{e^{-y}}{\sqrt{y}}\left[\left(1+\sqrt{2\Lambda y}\right)^{2N}+\left(1-\sqrt{2\Lambda y}\right)^{2N}\right], \tag{S18}$$

derived from the integral definition of $(2k-1)!!$ and the binomial theorem, Eq.(S17) becomes

$$\begin{aligned}
&Z_N(V,T,E_0,\Delta)\\
&\quad= \frac{(V/\lambda_0^3)^{2N}2^N}{2\pi(2N-1)!!}\int_0^\infty\int_0^\infty dx\,dy\,\frac{e^{-x}}{\sqrt{x}}\frac{e^{-y}}{\sqrt{y}}\,x^N\left[\left(1+\sqrt{y\delta\,(\eta+2x-1)\,/x}\right)^{2N}+(\sqrt{y}\to-\sqrt{y})\right]\\
&\quad= \frac{(V/\lambda_0^3)^{2N}2^N}{2\pi(2N-1)!!}\int_0^\infty\int_0^\infty dx\,dy\,\frac{e^{-x-y}}{\sqrt{xy}}\left[\left(\sqrt{x}+\sqrt{y\delta\,(\eta+2x-1)}\right)^{2N}+(\sqrt{y}\to-\sqrt{y})\right],
\end{aligned} \tag{S19}$$

where $(\sqrt{y}\to-\sqrt{y})$ stands in for the preceding term with $\sqrt{y}$ replaced with $-\sqrt{y}$. Next, using the identity

$$(2N-1)!! = \frac{2^N}{\sqrt{\pi}}\Gamma\left(N+1/2\right), \tag{S20}$$

gives the final form of the partition function.

4 Equilibrium Conditions for Non-Gendered System

In this section, we justify the conditions defining the Laplace's method approximation of the partition function and show that they result in a system of equations for $\langle k\rangle$ and $\langle m\rangle$, the average number of dimers and the average number of correct dimers, respectively.

In the main text, we made the approximation

$$\begin{aligned}
Z_N(V,T;E_0,\Delta) &= \int_0^\infty\int_0^\infty dx\,dy\,\exp\big[-\beta F_N(x,y;V,T,E_0,\Delta)\big]\\
&\simeq 2\pi\left(\det H\right)^{-1/2}\exp\big[-\beta F_N(\bar{x},\bar{y};V,T,E_0,\Delta)\big],
\end{aligned} \tag{S21}$$

where we defined

$$\beta F_N(x,y;V,T,E_0,\Delta) \equiv x+y+\frac{1}{2}\ln(xy)-\ln\left(\mathcal{M}_+^{2N}+\mathcal{M}_-^{2N}\right)+\beta F_0(N,V,T), \tag{S22}$$

with $\beta F_0(N,V,T)$ composed of terms that are independent of the variables x and y and of the parameters E_0 and Δ. In Eq.(S21), $\bar{x}$ and $\bar{y}$ are the critical points of $F_N(x,y;V,T,E_0,\Delta)$, defined by

$$\partial_i(\beta F_N)\Big|_{x=\bar{x},y=\bar{y}} = 0, \tag{S23}$$

for $i=x,y$, and H is the Hessian matrix with the elements

$$H_{ij} = \partial_i\partial_j(\beta F_N)\Big|_{x=\bar{x},y=\bar{y}}. \tag{S24}$$

For the validity of Eq.(S21), H must satisfy

$$\det H>0\,,\quad \mathrm{Tr}\,H>0. \tag{S25}$$

Eq.(S25) also ensures that the critical points defined by Eq.(24) are stable. We can compute the average number of dimers and the average number of correct dimers from the partition function via

$$\langle k\rangle = \frac{\partial}{\partial(\beta E_0)} \ln Z_N(V,T;E_0,\Delta), \qquad \langle m\rangle = \frac{\partial}{\partial(\beta\Delta)} \ln Z_N(V,T;E_0,\Delta). \tag{S26}$$

In Sec. 4.1, we will use the conditions Eq.(S23) along with the definitions in Eq.(S26) to calculate equilibrium constraints on $\langle k\rangle$ and $\langle m\rangle$. In Sec. 4.2, we will show the equilibria derived from these conditions satisfy Eq.(S25) and are indeed stable. Also, by computing the Hessian, we will show that the $\ln\det H$ contribution the Hessian could make to the free energy in Eq.(S22) is sub-leading in the large N limit because it is of the same order as the terms we drop in our derivation of the equilibrium conditions.

4.1 Computing Critical Points

Here we will derive the equilibrium conditions on $\langle k\rangle$ and $\langle m\rangle$ resulting from a $N \gg 1$ approximation of the partition function. We write the free energy Eq.(S22) slightly differently as

$$\beta F_N(x,y;V,T,E_0,\Delta) = x + y + (1/2 - N)\ln x + \frac{1}{2}\ln y - \ln\left(\mathcal{N}_+^{2N} + \mathcal{N}_-^{2N}\right) + \beta F_0(N,V,T), \tag{S27}$$

where

$$\mathcal{N}_\pm \equiv 1 \pm \delta^{1/2}\sqrt{y\,\Lambda(x;\beta\Delta)}, \tag{S28}$$

and

$$\Lambda(x;\beta\Delta) \equiv \frac{e^{\beta\Delta}-1}{x} + 2\,, \qquad \delta \equiv \frac{2\sqrt{2}\lambda_0^3}{V} e^{\beta E_0}. \tag{S29}$$

We can simplify Eq.(S27) by considering our presumed $N \gg 1$ limit. First we note that $(1+Q)^N + (1-Q)^N = (1+Q)^N + \phi_N$ where, if $Q > 0$, then $\phi_N \to 0$ as an inverse power of N for $N \to \infty$. Thus, Eq.(S27) can be written as

$$\beta F_N(x,y;V,T,E_0,\Delta) = x + y - N\ln x + \frac{1}{2}\ln y - 2N\ln\mathcal{N}_+ + \beta F_0(N,V,T) + \epsilon_N, \tag{S30}$$

where ϵ_N is the error term which includes all terms that are subleading in the $N \gg 1$ limit to the shown quantities. Now, using Eq.(S28) and Eq.(S30), we see that Eq.(S23) yield the equations

$$0 = \partial_x(\beta F_N)\Big|_{x=\bar{x},y=\bar{y}} = 1 - \frac{N}{\bar{x}} - \frac{N\delta^{1/2}\sqrt{\bar{y}/\Lambda(\bar{x};\beta\Delta)}}{1+\delta^{1/2}\sqrt{\bar{y}\Lambda(\bar{x};\beta\Delta)}} \cdot \left(-\frac{e^{\beta\Delta}-1}{\bar{x}^2}\right), \tag{S31}$$

$$0 = \partial_y(\beta F_N)\Big|_{x=\bar{x},y=\bar{y}} = 1 + \frac{1}{2\bar{y}} - \frac{N\delta^{1/2}\sqrt{\Lambda(\bar{x};\beta\Delta)/\bar{y}}}{1+\delta^{1/2}\sqrt{\bar{y}\Lambda(\bar{x};\beta\Delta)}}. \tag{S32}$$

From the definitions in Eq.(S26), we can express $\langle k\rangle$ and $\langle m\rangle$ in terms of $\bar{x}$ and $\bar{y}$:

$$\begin{aligned}\langle k\rangle &= \partial_{\beta E_0}\ln Z_N = -\partial_{\beta E_0}(\beta F_N)\Big|_{x=\bar{x},y=\bar{y}} \\ &= \frac{N\delta^{1/2}\sqrt{\bar{y}\Lambda(\bar{x};\beta\Delta)}}{1+\delta^{1/2}\sqrt{\bar{y}\Lambda(\bar{x};\beta\Delta)}}\end{aligned} \tag{S33}$$

$$\begin{aligned}\langle m\rangle &= \partial_{\beta\Delta}\ln Z_N = -\partial_{\beta\Delta}(\beta F_N)\Big|_{x=\bar{x},y=\bar{y}} \\ &= \frac{N\delta^{1/2}\sqrt{\bar{y}/\Lambda(\bar{x};\beta\Delta)}}{1+\delta^{1/2}\sqrt{\bar{y}\Lambda(\bar{x};\beta\Delta)}} \cdot \frac{e^{\beta\Delta}}{\bar{x}},\end{aligned} \tag{S34}$$

where we used Eq.(S31) and Eq.(S32) to set the coefficients of $\partial\bar{x}/\partial(\beta E_0)$ and $\partial\bar{y}/\partial(\beta E_0)$ (and similarly for the $\bar{x}$ and $\bar{y}$ derivatives with respect to $\beta\Delta$) to zero. To be explicit, we note that the second equalities in both Eq.(S33) and Eq.(S34) would be better expressed as approximations derived from Eq.(S21). However, for the analytical calculations of this system we will always be working in the $N \gg 1$ regime and we will take the free energy Eq.(S30) as the true free energy of the system.

From Eq.(S32), we find the condition

$$\bar{y} + 1/2 = \frac{N\delta^{1/2}\sqrt{\bar{y}\Lambda(\bar{x};\beta\Delta)}}{1+\delta^{1/2}\sqrt{\bar{y}\Lambda(\bar{x};\beta\Delta)}}, \tag{S35}$$

and with Eq.(S33), we obtain

$$\bar{y} + 1/2 = \langle k\rangle. \tag{S36}$$

Inverting Eq.(S35), we find

$$\delta\,\bar{y}\,\Lambda(\bar{x};\beta\Delta) = \frac{(\bar{y}+1/2)^2}{\left(N-(\bar{y}+1/2)\right)^2}, \tag{S37}$$

or, with Eq.(S36),

$$\delta\left(\langle k\rangle - 1/2\right)\Lambda(\bar{x};\beta\Delta) = \frac{\langle k\rangle^2}{\left(N-\langle k\rangle\right)^2}. \tag{S38}$$

We can further reduce this result by solving for $\Lambda(\bar{x};\beta\Delta)$ in terms of $\langle k\rangle$ and $\langle m\rangle$. Dividing Eq.(S33) by Eq.(S34), yields

$$\frac{\langle k\rangle}{\langle m\rangle} = \bar{x}\,\Lambda(\bar{x};\beta\Delta)e^{-\beta\Delta}, \tag{S39}$$

which when solved for $\bar{x}$, gives us

$$\bar{x} = \frac{1}{2}\left[1+\frac{\langle k\rangle - \langle m\rangle}{\langle m\rangle}e^{\beta\Delta}\right]. \tag{S40}$$

Substituting Eq.(S40) into Eq.(S39), then gives us

$$\Lambda(\bar{x};\beta\Delta) = \frac{2\langle k\rangle}{\langle k\rangle - \langle m\rangle(1-e^{-\beta\Delta})}. \tag{S41}$$

Returning to Eq.(S38), we obtain

$$2\delta\left(1-\frac{1}{2\langle k\rangle}\right) = \frac{\langle k\rangle - \langle m\rangle(1-e^{-\beta\Delta})}{\left(N-\langle k\rangle\right)^2}. \tag{S42}$$

which is the first equilibrium condition constraining $\langle k\rangle$ and $\langle m\rangle$. We will primarily be interested in temperature ranges at which $\langle k\rangle$ assumes a non-trivial value much larger than of $\mathcal{O}(1)$. Thus we can take $\langle k\rangle \gg 1$ leading to the result

$$\frac{4\sqrt{2}\lambda_0^3}{V}\,e^{\beta E_0} = \frac{\langle k\rangle - \langle m\rangle(1-e^{-\beta\Delta})}{\left(N-\langle k\rangle\right)^2} + \mathcal{O}(\langle k\rangle^{-1}) \tag{S43}$$

To find the second equilibrium condition, we note that Eq.(S31) can be written as

$$(N-\bar{x})\Lambda(\bar{x};\beta\Delta)\bar{x} = \langle k\rangle(e^{\beta\Delta}-1). \tag{S44}$$

Using Eq.(S39) and Eq.(S40), this result becomes

$$N - \frac{1}{2}\left[1+\frac{\langle k\rangle - \langle m\rangle}{\langle m\rangle}e^{\beta\Delta}\right] = \langle m\rangle(1-e^{-\beta\Delta}), \tag{S45}$$

or, with some rearranging,

$$\frac{e^{\beta\Delta}}{2} = \langle m\rangle \frac{N - \langle m\rangle(1 - e^{-\beta\Delta})}{\langle k\rangle - \langle m\rangle(1 - e^{-\beta\Delta})}, \tag{S46}$$

which is our second equilibrium condition. With the equilibrium conditions Eq.(S43) and Eq.(S46) established, we can now turn to showing that these equilibria define stable minima of the free energy.

4.2 Demonstrating Stability

To check whether the equilibrium conditions Eq.(S43) and Eq.(S46) define stable equilibria for this system, we need to compute the various elements of the Hessian matrix

$$H_{ij} = \partial_i\partial_j(\beta F_N)\Big|_{x=\bar{x},y=\bar{y}}, \tag{S47}$$

and ensure that the matrix is positive definite. By definition, a positive definite matrix is one with positive eigenvalues. For the 2×2 matrix considered here, this amounts to having a positive determinant and positive trace:

$$\mathrm{Tr}\, H > 0, \qquad \det H > 0. \tag{S48}$$

We will first compute the diagonal elements composing $\mathrm{Tr}\, H$. To compute $\partial_y^2(\beta F_N)|_{x=\bar{x},y=\bar{y}}$, we must compute the first and second-order y derivatives of the free energy as general functions. Given Eq.(S30), we obtain

$$\partial_y(\beta F_N) = 1 + \frac{1}{2y} - \frac{2N}{\mathcal{N}_+}\partial_y\mathcal{N}_+ \tag{S49}$$

$$\partial_y^2(\beta F_N) = -\frac{1}{2y^2} + 2N\left[\frac{(\partial_y\mathcal{N})_+^2}{\mathcal{N}_+^2} - \frac{\partial_y^2\mathcal{N}_+}{\mathcal{N}_+}\right]. \tag{S50}$$

From Eq.(S28), we have

$$\partial_y\mathcal{N}_+ = \frac{\delta^{1/2}}{2}\sqrt{\frac{\Lambda(x;\beta\Delta)}{y}}, \qquad \partial_y^2\mathcal{N}_+ = -\frac{\delta^{1/2}}{4}\sqrt{\frac{\Lambda(x;\beta\Delta)}{y^3}} = -\frac{1}{2y}\partial_y\mathcal{N}_+. \tag{S51}$$

Thus, Eq.(S50) becomes

$$\partial_y^2(\beta F_N) = -\frac{1}{2y^2} + 2N\left[\frac{(\partial_y\mathcal{N})_+^2}{\mathcal{N}_+^2} + \frac{1}{2y}\frac{\partial_y\mathcal{N}_+}{\mathcal{N}_+}\right]. \tag{S52}$$

Setting $x = \bar{x}$ and $y = \bar{y}$ in Eq.(S52) and noting that $\partial_y(\beta F_N) = 0$ at these values, we find

$$\partial_y^2(\beta F_N)\Big|_{x=\bar{x},y=\bar{y}} = \frac{1}{2N\bar{y}^2}\big[-N + (\bar{y}+1/2)^2 + N(\bar{y}+1/2)\big], \tag{S53}$$

where we used Eq.(S49) evaluated at $x = \bar{x}$ and $y = \bar{y}$. Considering the argument of the above expression, we find that it is positive for $\bar{y} > 1/2 + \mathcal{O}(N^{-1})$. In terms of our order parameter, this result translates into $\partial_y^2(\beta F_N)|_{x=\bar{x},y=\bar{y}}$ being positive for $\langle k\rangle > 1$ which is only violated when we are well-outside the range for non-trivial values of $\langle k\rangle$.

Next, computing $\partial_x^2(\beta F_N)$ given Eq.(S30), we obtain

$$\partial_x(\beta F_N) = 1 - \frac{N}{x} - \frac{2N}{\mathcal{N}_+}\partial_x\mathcal{N}_+ \tag{S54}$$

$$\partial_x^2(\beta F_N) = \frac{N}{x^2} + 2N\left[\frac{(\partial_x \mathcal{N})_+^2}{\mathcal{N}_+^2} - \frac{\partial_x^2 \mathcal{N}_+}{\mathcal{N}_+}\right], \tag{S55}$$

where

$$\partial_x \mathcal{N}_+ = \frac{\delta^{1/2}}{2}\sqrt{\frac{y}{\Lambda(x;\beta\Delta)}} \cdot \partial_x \Lambda(x;\beta\Delta), \tag{S56}$$

and

$$\partial_x^2 \mathcal{N}_+ = \frac{\partial_x \mathcal{N}_+}{\partial_x \Lambda(x;\beta\Delta)} \cdot \frac{1}{\Lambda(x;\beta\Delta)} \cdot \left[\Lambda(x;\beta\Delta)\,\partial_x^2\Lambda(x;\beta\Delta) - \frac{1}{2}\left(\partial_x\Lambda(x;\beta\Delta)\right)^2\right]. \tag{S57}$$

Using the definition of $\Lambda(x;\beta\Delta)$ (given in Eq.(S29)) in the quantity in the brackets above yields

$$\left[\Lambda(x;\beta\Delta)\,\partial_x^2\Lambda(x;\beta\Delta) - \frac{1}{2}\left(\partial_x\Lambda(x;\beta\Delta)\right)^2\right] = -\frac{\partial_x\Lambda(x;\beta\Delta)}{x}\left[\frac{3}{2}\Lambda(x;\beta\Delta) + 1\right]. \tag{S58}$$

Thus Eq.(S57) becomes

$$\partial_x^2 \mathcal{N}_+ = -\frac{\partial_x \mathcal{N}_+}{x}\left[\frac{3}{2} + \frac{1}{\Lambda(x;\beta\Delta)}\right]. \tag{S59}$$

Now, returning to Eq.(S55) we have

$$\partial_x^2(\beta F_N) = \frac{N}{x^2} + 2N\left(\frac{\partial_x \mathcal{N}_+}{\mathcal{N}_+}\right)^2 + 2N\frac{\partial_x \mathcal{N}_+}{\mathcal{N}_+}\left[\frac{3}{2x} + \frac{1}{x\Lambda(x;\beta\Delta)}\right]. \tag{S60}$$

Setting $x = \bar{x}$ and $y = \bar{y}$ in Eq.(S60) and noting that $\partial_x(\beta F_N) = 0$ at these values, we obtain

$$\partial_x^2(\beta F_N)\Big|_{x=\bar{x},y=\bar{y}} = \frac{1}{2N\bar{x}^2}\left[\bar{x}(\bar{x}+N) + \frac{2N(\bar{x}-N)}{\Lambda(\bar{x};\beta\Delta)}\right]. \tag{S61}$$

We can make further progress by expressing $\Lambda(\bar{x};\beta\Delta)$ in terms of $\bar{x}$ and $\bar{y}$. First, we note that Eq.(S40) and Eq.(S46) together yield

$$\bar{x} = N - \langle m\rangle(1 - e^{-\beta\Delta}), \tag{S62}$$

and inverting Eq.(S41) gives us

$$\frac{1}{\Lambda(\bar{x};\beta\Delta)} = \frac{1}{2}\left(1 - \frac{\langle m\rangle(1-e^{-\beta\Delta})}{\langle k\rangle}\right) = \frac{1}{2}\left(1 - \frac{N-\bar{x}}{\bar{y}+1/2}\right), \tag{S63}$$

where we used Eq.(S62) and Eq.(S36) in the final equality. Returning to Eq.(S61), we find

$$\partial_x^2(\beta F_N)\Big|_{x=\bar{x},y=\bar{y}} = \frac{1}{2N\bar{x}^2}\left[\bar{x}^2(\lambda+1) - 2N\bar{x}(\lambda-1) + N^2(\lambda-1)\right], \tag{S64}$$

where we defined

$$\lambda \equiv \frac{N}{\bar{y}+1/2}. \tag{S65}$$

Since $\bar{y}+1/2 = \langle k\rangle$ and $\langle k\rangle < N$, we have $\lambda > 1$ for non-zero temperature. For the function

$$f(z) = z^2(\lambda+1) - 2z(\lambda-1) + \lambda - 1 \tag{S66}$$

where $z \in \mathbb{R}^+$, it can be shown that the minimum satisfies

$$[f(z)]_{\min} = \frac{\lambda-1}{\lambda+1}. \tag{S67}$$

Thus, for $\lambda > 1$, we find that $f(z) > 0$. Therefore, Eq.(S64) is greater than zero for equilibrium values $\bar{x}$ and $\bar{y}$. With Eq.(S53) and Eq.(S64), we can thus conclude

$$\operatorname{Tr} H = \left[\partial_y^2(\beta F_N) + \partial_x^2(\beta F_N)\right]\Big|_{x=\bar{x},y=\bar{y}} > 0, \tag{S68}$$

for $\langle m \rangle$ and $\langle k \rangle$ constrained by Eq.(S43) and Eq.(S46).

Now, we compute the off-diagonal elements that make up, together with the diagonal elements, $\det H$. Taking the y-partial derivative of Eq.(S54), we have

$$\partial_y\partial_x(\beta F_N) = -2N\left[\frac{1}{\mathcal{N}_+}\partial_y\partial_x\mathcal{N}_+ - \frac{1}{\mathcal{N}_+^2}\partial_x\mathcal{N}_+\partial_y\mathcal{N}_+\right]. \tag{S69}$$

From Eq.(S28), we have that the mixed partial of $\mathcal{N}_+$ is

$$\begin{aligned}\partial_y\partial_x\mathcal{N}_+ &= \partial_y\left[\frac{\delta^{1/2}}{2}\sqrt{\frac{y}{\Lambda(x;\beta\Delta)}}\cdot\partial_x\Lambda(x;\beta\Delta)\right]\\ &= \frac{\delta^{1/2}}{4}\sqrt{\frac{1}{y\Lambda(x;\beta\Delta)}}\cdot\partial_x\Lambda(x;\beta\Delta) = \frac{1}{2y}\partial_x\mathcal{N}_+,\end{aligned} \tag{S70}$$

where we used Eq.(S56), in the final equality. Evaluating Eq.(S69) at $x=\bar{x}$ and $y=\bar{y}$ and using

$$\frac{1}{\mathcal{N}_+}\partial_y\mathcal{N}_+\Big|_{x=\bar{x},y=\bar{y}} = \frac{1}{2N}\left(1+\frac{1}{2\bar{y}}\right), \qquad \frac{1}{\mathcal{N}_+}\partial_x\mathcal{N}_+\Big|_{x=\bar{x},y=\bar{y}} = \frac{1}{2N}\left(1-\frac{N}{\bar{x}}\right), \tag{S71}$$

found from Eq.(S49), Eq.(S54), and the critical point condition, we obtain

$$\partial_y\partial_x(\beta F_N)\Big|_{x=\bar{x},y=\bar{y}} = \frac{(N-\bar{x})\big(N-\bar{y}-1/2\big)}{2N\bar{x}\bar{y}}. \tag{S72}$$

Before we compute the determinant, it will prove useful to express the $\bar{y}$ in Eq.(S53) and Eq.(S72) in terms of λ given in Eq.(S65). From Eq.(S65), we find

$$\begin{aligned}\partial_y^2(\beta F_N)\Big|_{x=\bar{x},y=\bar{y}} &= \frac{(\bar{y}+1/2)^2}{2N\bar{y}^2}\left(-\frac{N}{(\bar{y}+1/2)^2}+1+\frac{N}{\bar{y}+1/2}\right)\\ &= \frac{N}{2\lambda^2\bar{y}^2}\left(-\frac{\lambda^2}{N}+1+\lambda\right)\end{aligned} \tag{S73}$$

$$\begin{aligned}\partial_y\partial_x(\beta F_N)\Big|_{x=\bar{x},y=\bar{y}} &= \frac{1}{2N\bar{x}\bar{y}}(\bar{y}+1/2)(N-\bar{x})\left(\frac{N}{\bar{y}+1/2}-1\right)\\ &= \frac{1}{2\lambda\bar{x}\bar{y}}(N-\bar{x})(\lambda-1)\end{aligned} \tag{S74}$$

Finally, computing the determinant of the Hessian from Eq.(S64), Eq.(S73), and Eq.(S74), we thus find

$$\begin{aligned}\det H &= \left[\partial_y^2(\beta F_N)\partial_x^2(\beta F_N) - (\partial_y\partial_x(\beta F_N))^2\right]\Big|_{x=\bar{x},y=\bar{y}}\\ &= \frac{1}{4\lambda^2\bar{x}^2\bar{y}^2}\left[A_\lambda\bar{x}^2 - 2NB_\lambda\bar{x} + B_\lambda\right],\end{aligned} \tag{S75}$$

where

$$A_\lambda \equiv \lambda\left(4-\frac{\lambda(\lambda+1)}{N}\right) \tag{S76}$$

$$B_\lambda \equiv (\lambda - 1)\left(2 - \frac{\lambda^2}{N}\right). \tag{S77}$$

We want to show that Eq.(S75) is always positive. We will employ a method similar to that used in showing that $\partial_x^2(\beta F_N)|_{x=\bar{x},y=\bar{y}}$ is positive. For the function

$$g(z) = A_\lambda z^2 - 2B_\lambda z + B_\lambda, \tag{S78}$$

where $z \in \mathbb{R}^+$, it can be shown that the minimum is given by

$$[g(z)]_{\min} = B_\lambda\left(1 - \frac{B_\lambda}{A_\lambda}\right). \tag{S79}$$

From Eq.(S76), Eq.(S77), and the condition $1 < \langle\lambda\rangle < N$, we find that $B_\lambda < A_\lambda$ for all valid λ. From this inequality, we find

$$\frac{B_\lambda}{A_\lambda} < 1, \quad \text{if } B_\lambda > 0, \quad \text{and} \quad \frac{B_\lambda}{A_\lambda} > 1, \quad \text{if } B_\lambda < 0. \tag{S80}$$

Thus, we can conclude that Eq.(S79) is always positive for the entire domain of z and for valid values of λ. Considering Eq.(S75) we then have

$$\det H = \left[\partial_y^2(\beta F_N)\partial_x^2(\beta F_N) - \left(\partial_y\partial_x(\beta F_N)\right)^2\right]\Big|_{x=\bar{x},y=\bar{y}} > 0. \tag{S81}$$

With Eq.(S68) and Eq.(S81), we can conclude that the Hessian matrix is positive definite and thus that the derived equilibrium conditions Eq.(S43) and Eq.(S46) define stable equilibria of the free energy Eq.(21), and, moreover, that our Laplace's method approximation of the partition function Eq.(23) is valid.

Finally, in Eq.(S75), we see that $\ln\det H$ is on the order of a linear combination of $\ln\bar{x}$ and $\ln\bar{y}$. Given that we ultimately dropped such terms from our calculation of the equilibrium conditions Eq.(S43) and Eq.(S46), we now see that we were also justified in ignoring the $\ln\det H$ contributions to our free energy.

5 Simulation of Dimer System

The simulation results in Fig. 4 were obtained using the Metropolis-Hastings Monte Carlo algorithm. We defined the microstate of our system by two lists: One defining the particles that are monomers and the other defining the monomer-monomer pairs making up the dimers. For example, a $2N = 10$ particle system, could have a microstate defined by the monomer list $[1, 4, 6, 9]$ and the dimer list $[(3, 5), (2, 8), (7, 10)]$. The free energy of a microstate was given by

$$f(k, m) = -kE_0 - m\Delta - k\,k_BT\ln(V/\lambda_0^3) - (2N - 2k)k_BT\ln(2\sqrt{2}\,V/\lambda_0^3), \tag{S82}$$

for a system with k dimers of which m consisted of correct dimers.

To efficiently explore the state space of the system, we used three different types of transitions with unique probability weights for each one. In the following, N_m and N_d represent the lengths of the monomer and dimer lists, respectively, before the transition.

1. **Monomer Association:** Two randomly chosen monomers are removed from the monomer list, joined as a pair, and the pair is appended to the dimer list. Weight = $\binom{N_\mathrm{m}}{2}/(N_\mathrm{d}+1)$
   Example: mon = $[1, 3, 4, 5, 6, 9]$ and dim = $[(2, 8), (7, 10)] \to$ mon = $[1, 4, 6, 9]$ and dim = $[(3, 5), (2, 8), (7, 10)]$; Weight = 15/3.

2. **Dimer Dissociation:** One randomly chosen dimer is removed from the dimer list, and both of its elements are appended to the monomer lists. Weight = $N_\mathrm{d}/\binom{N_\mathrm{m}+2}{2}$

Example: mon = $[6, 9]$ and dim = $[(1, 4), (3, 5), (2, 8), (7, 10)] \rightarrow$ mon = $[2, 6, 8, 9]$ and dim = $[(1, 4), (3, 5), (7, 10)]$; Weight = 4/6.

3. **Dimer Cross-Over:** Two dimers are chosen randomly. One randomly chosen element from one dimer is switched with a randomly chosen element of the other dimer. Weight =1.

   Example: dim = $[(1, 4), (3, 5)(7, 10)] \rightarrow$ dim = $[(1, 10), (3, 5), (4, 10)]$]; Weight = 1.

The third type of transition is unphysical but is necessary to ensure that the system can quickly escape kinetic traps that lead to inefficient sampling of the state space.

For each simulation step, there was a 1/3 probability of selecting each transition type and the suggested step was accepted with log-probability

$$\ln p_{\text{accept}} = -(f_{\text{fin.}} - f_{\text{init.}})/k_B T + \ln(\text{Weight}), \tag{S83}$$

where f_{fin} and f_{init} are the final and initial free energies of the microstate defined according to Eq.(S82), and "(Weight)" is the ratio between the number of ways to make the forward transition and the number of ways to make the reverse transition. This weight was chosen for each transition type to ensure that detailed balance was maintained. For impossible transitions (e.g., monomer association for a microstate with no monomers), p_{accept} was set to zero.

At each temperature, the simulation was run for $30,000$ time steps, of which the last 600 were used to compute ensemble averages of $\langle k \rangle$ and $\langle m \rangle$. These simulations were repeated 50 times and each point in Fig. 4 represents the average $\langle k \rangle$ and $\langle m \rangle$ over these runs. IPython code for procedure is found in the *Supplementary Code*.

6 Temperature Changes in Parameter Space

In Fig. S1 we depict how the plots in Fig. 5 change as we change the value of $k_B T$. We note that since the regions are defined by temperature dependent boundaries, changing the temperature of a system represented by a point also changes the arrangement of the boundaries that surround the point.

References

[1] B. H. Margolius, "Avoiding your spouse at a bridge party," *Mathematics Magazine*, vol. 74, no. 1, pp. 33–41, 2001.

[2] C. Chuan-Chong and K. Khee-Meng, *Principles and techniques in combinatorics*, vol. 2. World Scientific, 1992.

[3] K. Dill and S. Bromberg, *Molecular driving forces: statistical thermodynamics in biology, chemistry, physics, and nanoscience*. Garland Science, 2012.

[4] K. P. Murphy, D. Xie, K. S. Thompson, L. M. Amzel, and E. Freire, "Entropy in biological binding processes: estimation of translational entropy loss," *Proteins: Structure, Function, and Bioinformatics*, vol. 18, no. 1, pp. 63–67, 1994.

[5] L. M. Amzel, "Loss of translational entropy in binding, folding, and catalysis," *Proteins: Structure, Function, and Bioinformatics*, vol. 28, no. 2, pp. 144–149, 1997.

[6] R. Milo and R. Phillips, *Cell biology by the numbers*. Garland Science, 2015.

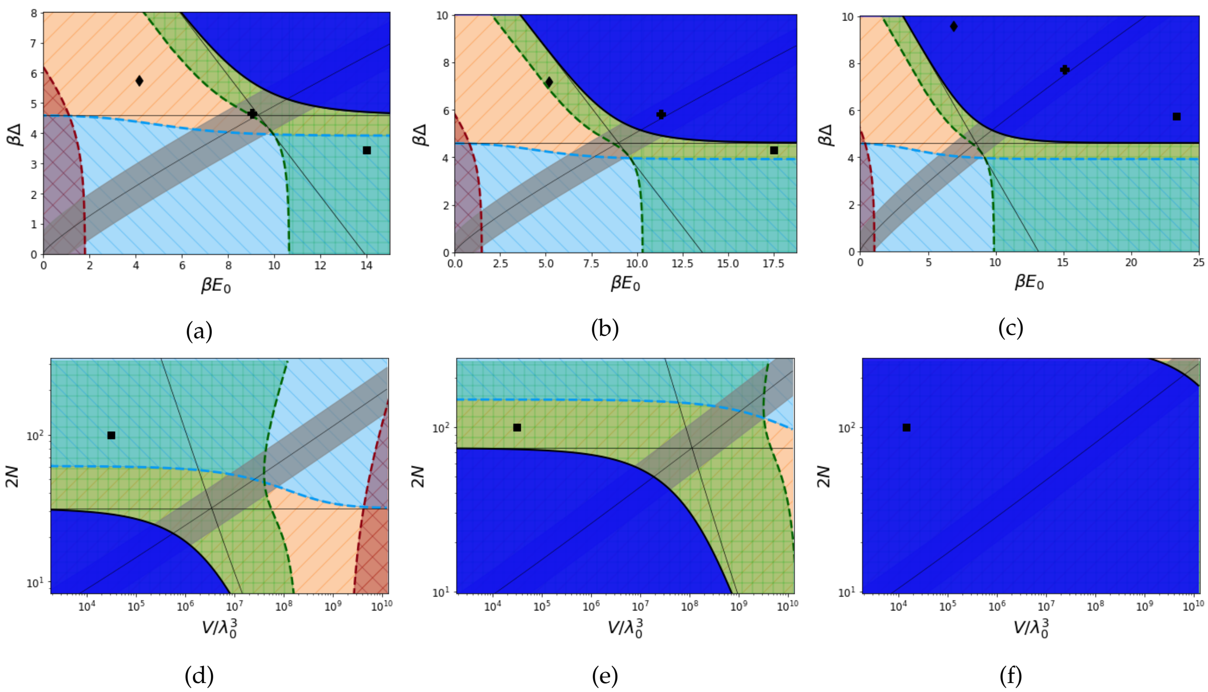

Figure S1: (a), (b), and (c): Plots of the points in Fig. 5a in the main text as we lower the system temperature: (a) depicts $k_BT = 1.0$ (i.e., the same temperature as the original figure), (b) depicts $k_BT = 0.8$, and (c) depicts $k_BT = 0.6$. Consistent with the simulation plots in Fig. 4, at $k_BT = 0.6$ all the systems are in the "fully correct assembly" regime. (d), (e), and (f): Plots of the point in Fig. 5b as we lower the system temperature: (d) depicts $k_BT = 1.0$ (i.e., the same temperature as the original figure), (e) depicts $k_BT = 0.8$, and (f) depicts $k_BT = 0.6$. Consistent with the simulation plots in Fig. 4(c), at $k_BT = 0.6$ the systems is in the "fully correct assembly" regime.