Free energy of ligand-receptor systems forming multimeric complexes

Lorenzo Di Michele,[1] Stephan J. Bachmann,[2] Lucia Parolini,[1] and Bortolo M. Mognetti[2, a)]
[1)] *Biological and Soft Systems, Cavendish Laboratory, University of Cambridge, JJ Thomson Avenue, Cambridge, CB3 0HE, United Kingdom*
[2)] *Interdisciplinary Center for Nonlinear Phenomena and Complex Systems & Service de Physique des Systèmes Complexes et Mécanique Statistique, Université libre de Bruxelles (ULB), Campus Plaine, CP 231, Blvd du Triomphe, B-1050 Brussels, Belgium.*

Ligand-receptor interactions are ubiquitous in biology and have become popular in materials in view of their applications to programmable self-assembly. Although, complex functionalities often emerge from the simultaneous interaction of more than just two linker molecules, state of art theoretical frameworks enable the calculation of the free energy only in systems featuring one-to-one ligand/receptor binding. In this communication we derive a general formula to calculate the free energy of systems featuring simultaneous direct interaction between an arbitrary number of linkers. To exemplify the potential and generality of our approach we apply it to the systems recently introduced by Parolini *et al.* [*ACS Nano* **10**, 2392 (2016)] and Halverson *et al.* [*J. Chem. Phys.* **144**, 094903 (2016)], both featuring functionalized Brownian particles interacting via three-linker complexes.

The quantitative understanding of the ligand-receptor interactions is receiving much attention in view of the key role played in biology and their applications to the self-assembly of composite materials.
Biological cells respond to the presence of specific molecules via cell-surface receptors. Examples include *toll-like receptors*, triggering immune response to bacterial and viral activity[1], and *receptor tyrosine kinases*, involved in the regulation of several physiological processes[2]. In order for the signals to be transduced across the cell membrane, the presence of the ligands typically triggers dimerization or oligomerization of the receptors, through interactions that involve multiple molecules.
Functionalizing Brownian units with specific linkers, often made of synthetic DNA molecules, is a powerful tool to engineer the structure and response of self-assembled soft materials[3–11]. Many functionalization schemes rely on one-to-one ligand-receptor interactions, but recently designs featuring multi-linker complexation have been proposed to extend the accessible range of functionalities[9,12–15]. In particular, Parolini *et al.*[13] adopted three-linker complexes enabling toehold-mediated strand exchange reactions[16] to control aggregation kinetics of lipid vesicles coated with DNA linkers. Halverson *et al.*[12] also proposed the use of three-linker DNA complexes to program a cooperative behavior between functionalized particles, which could allow to control the sequence of binding events in the self-assembly. Recently, Angioletti-Uberti *et al.*[17] proposed an analytical expression for the free energy of systems featuring one-to-one ligand-receptor interactions that overcame some limitations of earlier approaches[18,19]. In this Communication we provide a more general framework to calculate the free energy of systems including multimeric complexes featuring an arbitrary number of ligand/receptors (see Fig. 1). We consider "particles", e.g. biological cells or artificial colloidal units, functionalized by surface ligands/receptors ("linkers" or "molecules"). We assume that linkers can freely diffuse on the surface of the particles. An extension to immobile linkers can be derived following Ref.[20]. Bonds can either involve linkers tethered to the same particle or to different particles. Excluded volume interactions between the molecules are neglected. Our results are exact in the limit of many linkers per particle[21,22]. We envisage applications of our theory to the association of more complex molecules like DNA tiles[23–25] or virial caspids[26].
In Sec. I we derive our theory while in Sec. II we test it on the system introduced in Ref.[12], calculating the interaction free energy between particles and quantitatively justifying the postulated cooperative behavior. In Sec. III we examine the suspensions of DNA-functionalized vesicles of Ref.[13], discussing the thermodynamic ground state in relation to the kinetic behaviour characterized in the original publication.

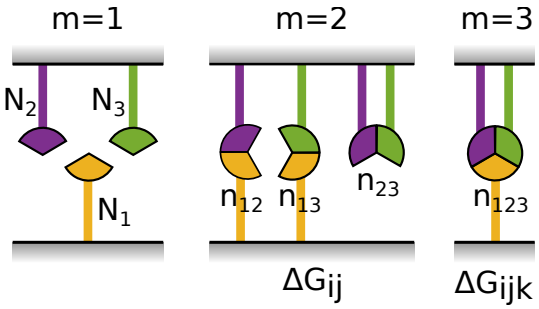

FIG. 1. A system of three families of linkers binding in pairs and three-molecule multimeric complexes

I. FREE ENERGY CALCULATION

We consider c families of different linkers, each with a number of units N_i $(i = 1, \cdots, c)$. The linkers can re-

[a)] Electronic mail: bmognett@ulb.ac.be

versibly associate into complexes of m units. For clarity we only consider complexes that never feature more than a single linker of each family ($1 \le m \le c$, see Fig. 1 where $c = 3$). In Sec. S1 of the supplemental material (SM)[27] we show that relaxing this assumption does not change our main result (Eq. 7). The state of the system is described by the number $n_{i_1,i_2,\cdots i_m}$ of all the possible complexes made by m linkers of type i_1, i_2, $\cdots$ i_m, with $i_1 < i_2 < \cdots i_m$ and $2 \le m \le c$.

We start by deriving an expression for the partition function Z of the system as the weighed sum over all the possible realizations of $n_{i_1,i_2,\cdots i_m}$. First we calculate the contribution of two-linker complexes $\{n_{i_1,i_2}\}$ to Z, then we deplete the total number of linkers of each family N_i by the number of those involved in two-linker complexes and calculate the contribution from complexes with three molecules $\{n_{i_1,i_2,i_3}\}$. This procedure is repeated recursively. When calculating the contribution of complexes with $m+1$ linkers, N_i has been reduced to $N_i^{(m)}$ that is given by

$$n_i^{(m)} = \sum_{i_2<\cdots<i_m} n_{\tau[i,i_2,\cdots i_m]} \quad N_i^{(m)} = N_i - \sum_{\ell=2}^{m} n_i^{(\ell)} \quad (1)$$

where $n_i^{(\ell)}$ is the total number of linkers of type i involved in complexes of size ℓ, and τ is the operator that orders m indices. $N_i^{(c)}$ is the number of linkers of type i that are free, and will be also indicated by n_i below. The partition function is then expressed as

$$Z = \sum_{\{n_{i_1,i_2}\}} Z^{(2)} \sum_{\{n_{i_1,i_2,i_3}\}} Z^{(3)} \cdots \sum_{\{n_{i_1,i_2,\cdots i_c}\}} Z^{(c)}, \quad (2)$$

where the curly brackets indicate the ensemble of all the complexes formed by a given number of linkers. Note that in Eq. 2 $Z^{(\ell)}$ is a function of $\{N_i^{(\ell-1)}\}_{1\le i\le c}$ and, as a consequence of Eq. 1, of the number of complexes with $m \le \ell$.

Defining $\Delta G_{i_1,\cdots i_m}$ as the free energy associated to the formation of a $i_1 \cdots i_m$ complex[28], we can define the contribution to the partition function from all the complexes of size m as

$$Z^{(m)} = \Omega^{(m)} \Big(\{N_i^{(m-1)}\}; \{n_{i_1,i_2,\cdots,i_m}\} \Big) \exp\Big[- \sum_{i_1<i_2\cdots<i_m} n_{i_1,i_2\cdots,i_m} \beta \Delta G_{i_1,i_2\cdots,i_m} \Big], \quad (3)$$

where $\beta = 1/(k_{\rm B}T)$ and $\Omega^{(m)}$ accounts for the combinatorial factors. The latter can be written as

$$\Omega^{(m)} = \prod_{i=1}^{c} \frac{N_i^{(m-1)}!}{N_i^{(m)}!} \prod_{i_1<i_2<\cdots<i_m} \frac{1}{n_{i_1,i_2,\cdots,i_m}!}, \quad (4)$$

where the first product counts the number of ways one can choose the molecules belonging to the complexes $\{n_{i_1,i_2,\cdots,i_m}\}$ starting from $\{N_i^{(m-1)}\}$ free linkers, while the second term accounts for the number of independent ways to build such a set of complexes. Using Eq. 4 and Eq. 3 into Eq. 2, we can calculate the partition function and the free energy F of the system

$$Z = e^{-\beta F} = \sum_{\{\{n_{i_1,i_2,\cdots,i_\ell}\}\}} \exp[-\beta \mathcal{A}(\{\{n_{i_1,i_2,\cdots,i_\ell}\}\})] \quad (5)$$
$$= \sum_{\{\{n_{i_1,i_2,\cdots,i_\ell}\}\}} \prod_{i=1}^{c} \frac{N_i!}{n_i!} \prod_{m=2}^{c} \prod_{i_1<i_2<\cdots<i_m} \frac{1}{n_{i_1,i_2,\cdots,i_m}!} \exp\Big[- \sum_{i_1<i_2\cdots<i_m} n_{i_1,i_2\cdots,i_m} \beta \Delta G_{i_1,i_2\cdots,i_m} \Big],$$

where the double curly brackets $\{\{\dots\}\}$ indicate the ensemble of complexes of arbitrary size, and $\mathcal{A}$ is a functional introduced for later convenience.

In the limit of large N_i we can simplify Eq. 5 using a saddle point approximation. In particular the stationary point of $\mathcal{A}$, given by $\delta\mathcal{A}/\delta n_{i_1,\cdots i_m} = 0$, identifies the average number of complexes $\overline{n}_{i_1 i_2 \cdots i_m} = \langle n_{i_1 i_2 \cdots i_m} \rangle$. The stationary conditions for the functional $\mathcal{A}$ as defined by Eq. 5 become

$$\overline{n}_{i_1 i_2 \cdots i_m} = \overline{n}_{i_1} \overline{n}_{i_2} \cdots \overline{n}_{i_m} \exp[-\beta \Delta G_{i_1,\cdots i_m}] \ . \quad (6)$$

Note that Eq. 6 are the relations for chemical equilibrium expressed in terms of the total number of molecules. When considering tethered linkers (Fig. 1) the complexation free energy $\Delta G_{i_1,\cdots i_m}$[28] also includes rotational and translational entropic costs controlled by the length of the spacers and by the size of the particles[7].

Using Eq. 6 into Eq. 5 to express $\Delta G_{i_1,\cdots i_m}$ as a function of equilibrium number of complexes, we can evaluate the free energy of the system as $F = \mathcal{A}(\{\{\overline{n}_{i_1,i_2,\cdots,i_\ell}\}\})$. By considering only the dominant term in the second line of Eq. 5, and using Stirling's approximation we find

$$\begin{aligned}
\beta F &= \sum_{i=1}^{c} \overline{n}_i \log \overline{n}_i - N_i \log N_i - \overline{n}_i + N_i \\
&\quad + \sum_{m=2}^{c} \sum_{i_1<i_2\cdots<i_m} \overline{n}_{i_1,i_2,\cdots,i_m} \left(\log \overline{n}_{i_1} + \cdots \log \overline{n}_{i_m} - 1 \right) \\
&= \sum_{i=1}^{c} N_i \log \frac{\overline{n}_i}{N_i} - \overline{n}_i + N_i - \sum_{m=2}^{c} \sum_{i_1<i_2\cdots<i_m} \overline{n}_{i_1,i_2\cdots,i_m} \\
&= \sum_{i=1}^{c} N_i \log \frac{\overline{n}_i}{N_i} + \sum_{i=1}^{c} \sum_{m=2}^{c} \sum_{i_2<i_3\cdots<i_m} \overline{n}_{\tau[i,i_2,i_3,\cdots,i_m]} \\
&\quad - \sum_{m=2}^{c} \sum_{i_1<i_2\cdots<i_m} \overline{n}_{i_1,i_2,\cdots,i_m} \ ,
\end{aligned}$$

where Eq. 1 has been used in the second equality to factorize the terms $\log n_i$, and in the third equality to express $N_i - n_i$ in terms of higher order complexes. Finally we obtain the main result of this work

$$\beta F = \sum_{i=1}^{c} N_i \log \frac{\overline{n}_i}{N_i} + \sum_{m=2}^{c} (m-1) \sum_{i_1<i_2\cdots<i_m} \overline{n}_{i_1,i_2,\cdots,i_m}. \quad (7)$$

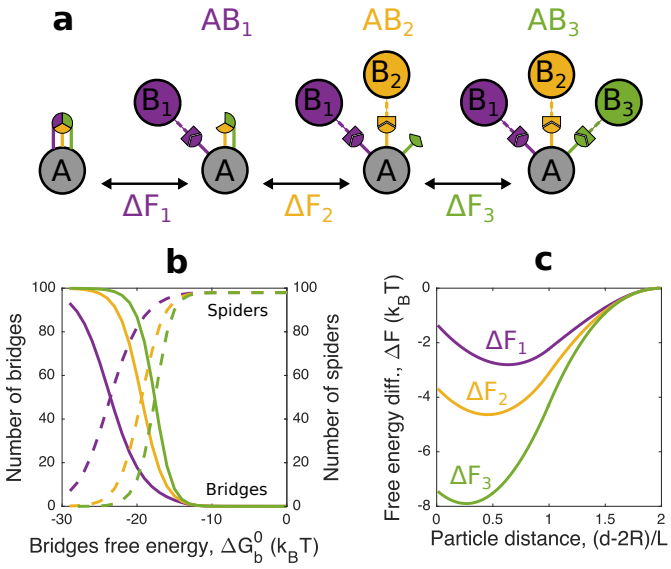

FIG. 2. Cooperative binding scheme of Ref.[12]. **a**, For isolated A particles most linkers form spiders. For B_1 to bind, stable spiders need to break to form bridges. Loops are left on A. The binding of B_2 is more favourable, as loops are less stable than spiders. Binding the third particle is most favorable, as no intra-particle complexes need to open on A. **b**, Number of spiders and bridges in AB_n complexes, color-coded following panel **a**. **c**, Free energy difference between the complexes shown in panel **a** as a function of inter-particle distance for $\Delta G_b^0 = -14 k_B T$.

Being written in term of the equilibrium number of complexes, Eq. 7 cannot be used to sketch free-energy landscapes[24,25], however it is applicable to calculate effective interactions between functionalized objects as demonstrated in the next two sections. Note that in order to guarantee the extensivity of the functional $\mathcal{A}$, $\Delta G_{i_1,\dots i_m} - (m-1)\log N$ (with $N_i \sim N$) should be kept fixed when taking the $N \to \infty$ limit (see also Eq. 6). For the case of one-to-one interactions ($m = 2$), Eq. 7 reduces to the result of Ref.[17].

II. BINDING COOPERATIVITY IN DNA-FUNCTIONALIZED PARTICLES

As a first example we examine the cooperative self-assembly scheme recently proposed by Halverson and Tkachenko[12], based on the possibility of forming three-linker complexes, dubbed spiders. As shown in Fig. 2**a**, we consider particles of type A functionalized by $3N$ mobile DNA linkers equally distributed among three families, each carrying different single-stranded DNA sequences or *sticky-ends*, labelled as α_i, $i = 1, 2, 3$. Such sticky-ends can hybridize to form three different families of intra-particle *loops* (ℓ_i), involving two out of three types of linkers, or spiders (s), involving all three types (see Fig. 2**a**). We then consider three types of particles B_i, $i = 1, 2, 3$, each functionalized by N identical linkers carrying a sticky-end sequence $\overline{\alpha}_i$ complementary to α_i. Linkers on particles B_i can form inter-particle *bridges* b_i, with particles A. In the following we consider linkers constituted by double stranded DNA spacers of length $L = 10\,\text{nm}$ and point-like sticky ends[21], rigid particles of radius $R = 10\,L$[21], and $N = 100$. See SM Sec. S2 and to Ref.[21] for details. Below we calculate the free energy $F(AB_n)$ of clusters made by a single A particle and a variable number n of B particles taken as in Fig. 2. We demonstrate a cooperative effect by which the free energy gain from binding the n-th B particle $\Delta F_n = F(AB_n) - F(AB_{n-1})$ is higher than the gain from binding the $(n-1)$-th one, for $n = 2, 3$. This is due to the necessity of breaking spider and loop complexes formed on the A particle for the 1st or the 2nd B particles to bind. Our theory allows to calculate the free energy gain for binding the 1st, the 2nd and the 3rd B particles, chosen as a model parameters in Ref.[12]. The number of complexes at equilibrium are given by[21]

$$\overline{n}_{\ell_k} = \overline{n}_{\alpha_i}\overline{n}_{\alpha_j}\left[\frac{e^{-\beta\Delta G_\ell^0}}{\rho_\ominus v_A}\right] \quad (k \neq i,j \text{ AND } i \neq j) \tag{8}$$

$$\overline{n}_s = \overline{n}_{\alpha_1}\overline{n}_{\alpha_2}\overline{n}_{\alpha_3}\left[\frac{e^{-\beta\Delta G_s^0}}{(\rho_\ominus v_A)^2}\right] \tag{9}$$

$$\overline{n}_{b_i} = \overline{n}_{\alpha_i}\overline{n}_{\bar{\alpha}_i}\left[\frac{\epsilon_i v_{AB} e^{-\beta\Delta G_b^0}}{\rho_\ominus v_A v_B}\right] \tag{10}$$

where ΔG_ℓ^0, ΔG_s^0 and ΔG_b^0 are the hybridization free energies of the sticky-ends associated to loop, spider, and bridge formation respectively, $\rho_\ominus$=1M is the standard concentration, and $v_{A/B/AB}$ are volume factors reported in SM Sec. S2 that quantify the configurational entropic costs of binding mobile tethers (Refs.[7,21] and SM Sec. I). Note that by using the generic notation of Sec. I we would have $\overline{n}_{\ell_k} \equiv \overline{n}_{\alpha_i\alpha_j}$, $\overline{n}_s \equiv \overline{n}_{\alpha_1\alpha_2\alpha_3}$, and $\overline{n}_{b_i} \equiv \overline{n}_{\alpha_i\bar{\alpha}_i}$. Different types of loops and bridges are assumed to have the same hybridization free energy. In Eq. 10 $\epsilon_i = 0, 1$ specifies if B_i is bound or not to A. In particular $n = \sum_i \epsilon_i$ indicates the valence of particle A. Equations 8, 9, and 10 are then closed by the conditions $N = \overline{n}_{\alpha_i} + 2\overline{n}_\ell + \overline{n}_s + \overline{n}_{b_i}$ and $N = \overline{n}_{\alpha_i} + \overline{n}_{b_i}$.

First we consider an isolated A particle and calculate the number of loop and spider complexes as a function of ΔG_ℓ^0, choosing $\Delta G_s^0 = 3\Delta G_\ell^0$. As shown in Fig. S1 of the SM, when $\Delta G_\ell^0 = -10\,k_B T$ only spiders are present on A. We fix ΔG_ℓ^0 to this value as a reasonable guess to maximize the cooperative behaviour. We then consider particle clusters AB_n (with $n = 0, 1, 2, 3$), with distances between the centers of A and B particles equal to $d = 2R + L$, and calculate the number of bridges n_{b_i} as a function of ΔG_b^0 (see Fig. 2**b**). We find that bridges form at higher values of ΔG_b^0 when n is higher. Finally we use Eq. 7 (contextualized to this system in SM Eq. S14) to calculate the free energy of the system including the repulsive part of the potential calculated accounting for the entropic compression of the DNA strands between the particles (see SM Eq. S2). We consider clusters in which all of the B particles are at the same distance d

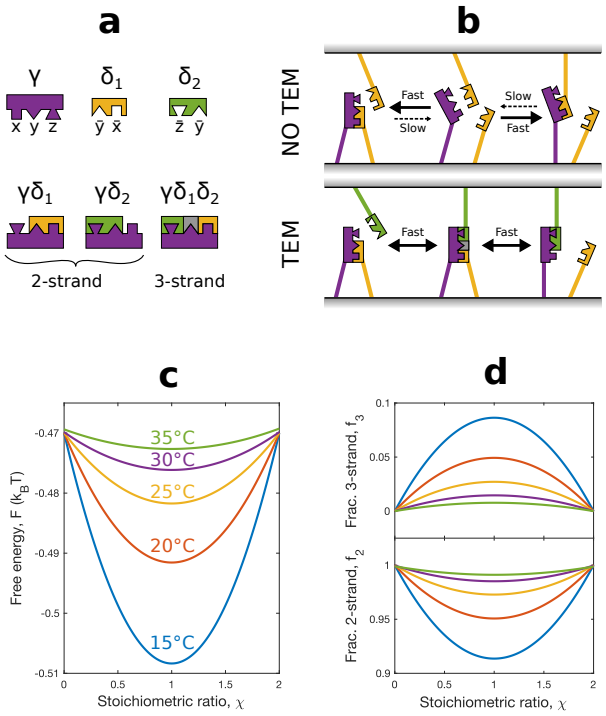

FIG. 3. Toehold-exchange mechanism of Ref[13]. **a**, We consider three sticky ends γ, δ_1 and δ_2 that can form two- or three-strand complexes. **b**, If only one between δ_1 and δ_2 are present ($\chi = 0, 2$), the formation of an inter-particle bridge requires the thermal breakup of an intra-particle bridge, making aggregation slow (Top). If both δ_1 and δ_2 are present, TEM catalyzes loop-to-bridge swap. **c**, Despite the difference in kinetic behavior, the free energy of the system (*per* γ strand) as calculated within our framework depends on χ only weakly regardless of temperature. **d**, Fraction of two-strand (f_2) and three-strand (f_3) complexes as calculated by our theory at different temperatures.

from the A-particle, and for which B-particles do not interact with each other (see Fig. 2**a**). Figure 2**c** shows the free-energy change associated to the binding of a single B particle to a cluster as a function of d. As expected, the free-energy gain obtained when adding the second B particle is higher than that obtained by binding the first, and the gain achieved upon adding the third particle is significantly higher than both the former.

We note that kinetic bottlenecks associated to the opening of the stable spider and loop complexes are likely to slow down self-assembly. Incidentally, strand-displacement strategies[16] similar to those discussed in the next section and in Ref.[13] can speed up relaxation.

III. INTERACTION FREE ENERGY IN THE PRESENCE OF TOEHOLD-EXCHANGE-MECHANISM

As a second example we examine the system studied experimentally in Ref.[13]. Let us consider a suspension of identical micron-size lipid vesicles, functionalized by three families of mobile DNA linkers with sticky ends here labelled as γ, δ_1, and δ_2. As shown in Fig. 3**a**, sticky end γ is made of three domains of equal length, x, y, and z. Sticky ends δ_1 have two domains $\bar{x}$ and $\bar{y}$, complementary to x and y, whereas δ_2 features domains $\bar{y}$ and $\bar{z}$. Linker γ can bind to δ_1 and δ_2, with comparable hybridization free energy. A three-linker $\gamma\delta_1\delta_2$ complex is also possible, where δ_1 and δ_2 bind to domains x and z respectively, and compete to occupy domain y. δ_1 does not bind to δ_2. Two- and three-linker complexes can form either among linkers tethered to the same vesicles (loop-like) or between different vesicles (bridge-like). At sufficiently high temperature all of the linkers are unbound. If the suspension is quenched to low temperature, the formation of intra-vesicle loop-like complexes is kinetically favored over the formation of bridges, effectively sequestrating all of the available γ linkers. The aggregation of the liposomes, mediated by the formation of inter-vesicle bridges, is therefore limited by the opening of the intra-vesicle loops, which seldom occurs at low temperatures (Fig. 3**b**, Top). Through a *Toehold-Exchange Mechanism* (TEM)[16], the formation of three-strand complexes mediates the swap between stable loops and stable bridges without the need for thermal denaturation. In particular, the toehold domain z (x) causes a free δ_2 (δ_1) linker to transiently bind to an existing $\gamma\delta_1$ ($\gamma\delta_2$) bond, facilitating the reaction $\gamma\delta_1 + \delta_2 \leftrightharpoons \delta_1 + \gamma\delta_2$ (Fig. 3**b**, Bottom). We indicate with $3N$ the total number of linkers per vesicle, N of which are of type γ, $N\chi$ of type δ_1, and $N(2-\chi)$ of type δ_2. The parameter $\chi \in [0, 2]$ controls the stoichiometric ratio between δ_1 and δ_2 and thereby the effectiveness of the TEM process. For $\chi = 0$ or 2 three-strand complexes are not possible and the bridge formation and aggregation kinetics are dominated by the slow opening of formed loops. For $\chi = 1$, TEM is most effective and aggregation kinetics is found to speed up by more than one orderer of magnitude at $T = 15°\mathrm{C}$[13].

We use our framework to calculate the free energy of the system, and demonstrate that, despite the large effect on aggregation kinetics, changing χ has little consequences on the thermodynamic ground state of the system. The DNA tethers are again modelled as freely pivoting rigid rods of length $L = 10$ nm, with freely diffusing tethering points and point-like sticky ends. For simplicity we model two interacting vesicles as flat planes of area $A = 0.5\mu\mathrm{m}^2$ kept at a distance of $h = 1.4L$ from each other. We chose $N = 360$. Hybridization free-energies between the sticky ends are taken from Ref.[13].

Explicit expression for the equilibrium distributions of all the possible complexes are shown in the SM Eqs. S15-17. In the SM (Eqs. S21, S22) we provide the expression for the interaction free energy between two vesicles (*per* γ strand), shown as a function of χ and T in Fig. 3**c**. We observe that regardless of temperature, the free energy decreases by less than 10% when going from $\chi = 0, 2$ to $\chi = 1$, supporting the claim that with the architecture proposed in Ref.[13] aggregation kinetics can be substan-

tially changed with little consequences on the thermodynamic ground state. The weak dependence of the overall free energy on χ is a direct consequence of the small number of three-strand complexes, always involving less than 10% of all γ linkers, as demonstrated in Fig. 3**d**.

IV. CONCLUSIONS

We provide an analytical expression for the free energy of systems of ligand/receptors that can form complexes featuring an arbitrary number of molecules. Our framework can be applied to biologically relevant situations, where cell-surface receptors form trimers or oligomers, or to suspensions of colloidal particles functionalized by synthetic DNA ligands: an increasingly popular strategy to achieve controlled self-assembly of complex soft materials. To exemplify the versatility of our approach, we re-examine the artificial systems recently proposed by Halverson *et al.*[12] and Parolini *et al.*[13], both featuring DNA-functionalized Brownian particles interacting trough the formation of three-linker complexes. For the former, we are able to quantify the cooperative effects in the interaction free energy between the particles, taken as model parameters in the original publication. For the system of Parolini *et al.* we study the interaction free energy between vesicles with different linker stoichiometry. Our theory demonstrates that despite the substantial effect on aggregation kinetics observed experimentally, coating stoichiometry has a comparatively small effect of the thermodynamic ground state of the suspension.

Acknowledgements. LDM and LP acknowledge support from the EPSRC Programme Grant CAPITALS number EP/J017566/1. LDM acknowledges support from the Oppenheimer Fund and Emmanuel College Cambridge. SB and BMM are supported by the Université Libre de Bruxelles (ULB). The python programs used to calculate the results presented here are available at `http://dx.doi.org/10.5281/zenodo.47204`.

[1] N. J. Gay, M. F. Symmons, M. Gangloff, and C. E. Bryant, Nat. Rev. Immunol. **14**, 546 (2014).
[2] M. A. Lemmon and J. Schlessinger, Cell **141**, 1117 (2010).
[3] C. A. Mirkin, R. C. Letsinger, R. C. Mucic, and J. J. Storhoff, Nature **382**, 607 (1996).
[4] A. P. Alivisatos, K. P. Johnsson, X. Peng, T. E. Wilson, C. J. Loweth, M. P. Bruchez, and P. G. Schultz, Nature **382**, 609 (1996).
[5] R. Macfarlane, B. Lee, M. Jones, N. Harris, G. Schatz, and C. A. Mirkin, Science **334**, 204 (2011).
[6] L. Di Michele and E. Eiser, Phys. Chem. Chem. Phys. **15**, 3115 (2013).
[7] L. Parolini, B. M. Mognetti, J. Kotar, E. Eiser, P. Cicuta, and L. Di Michele, Nat. Commun. **6**, 5948 (2015).
[8] S. Angioletti-Uberti, B. Mognetti, and D. Frenkel, Nat. Mater. **11**, 518 (2012).
[9] W. B. Rogers and V. N. Manoharan, Science **347**, 639 (2015).
[10] Y. Wang, Y. Wang, D. R., V. N. Manoharan, L. Feng, A. D. Hollingsworth, M. Weck, and D. J. Pine, Nature **491**, 51 (2012).
[11] Y. Wang, Y. Wang, X. Zheng, É. Ducrot, J. S. Yodh, M. Weck, and D. J. Pine, Nature Comm. **6** (2015).
[12] J. D. Halverson and A. V. Tkachenko, J. Chem. Phys. **144**, 094903 (2016).
[13] L. Parolini, J. Kotar, L. Di Michele, and B. M. Mognetti, ACS Nano **10**, 2392 (2016).
[14] F. Romano and F. Sciortino, Phys. Rev. Lett. **114**, 078104 (2015).
[15] J. T. McGinley, I. Jenkins, T. Sinno, and J. C. Crocker, Soft Matter **9**, 9119 (2013).
[16] D. Y. Zhang and E. Winfree, J. Am. Chem. Soc. **131**, 17303 (2009).
[17] S. Angioletti-Uberti, P. Varilly, B. Mognetti, A. Tkachenko, and D. Frenkel, J. Chem. Phys. **138**, 021102 (2013).
[18] R. Dreyfus, M. E. Leunissen, R. Sha, A. V. Tkachenko, N. C. Seeman, D. J. Pine, and P. M. Chaikin, Phys. Rev. Lett. **102**, 048301 (2009).
[19] S. Angioletti-Uberti, B. M. Mognetti, and D. Frenkel, Phys. Chem. Chem. Phys. **18**, 6373 (2016).
[20] P. Varilly, S. Angioletti-Uberti, B. Mognetti, and D. Frenkel, J. Chem. Phys. **137**, 094108 (2012).
[21] S. Angioletti-Uberti, P. Varilly, B. M. Mognetti, and D. Frenkel, Phys. Rev. Lett. **113**, 128303 (2014).
[22] H. Xu and D. E. Shaw, Biophys. J. **110**, 218 (2016).
[23] F. Dannenberg, K. E. Dunn, J. Bath, M. Kwiatkowska, A. J. Turberfield, and T. E. Ouldridge, J. Chem. Phys. **143**, 165102 (2015).
[24] W. M. Jacobs and D. Frenkel, J. Am. Chem. Soc. **138**, 2457 (2016).
[25] J. Fern, J. Lu, and R. Schulman, ACS nano **10**, 1836 (2016).
[26] E. C. Dykeman, P. G. Stockley, and R. Twarock, Proc. Natl. Acad. Sci. USA **111**, 5361 (2014).
[27] See supplemental material at [URL will be inserted by AIP] for a more general derivation of Eq. 7 and for additional details of the examples studied in Sec. II and Sec. III.
[28] R. M. Dirks, J. S. Bois, J. M. Schaeffer, E. Winfree, and N. A. Pierce, Siam Rev. **49**, 65 (2007).

SUPPLEMENTAL MATERIAL: Free energy of lingand-receptor systems forming multimeric complexes

Lorenzo Di Michele,[1] Stephan J. Bachmann,[2] Lucia Parolini,[1] and Bortolo M. Mognetti[2, a)]

[1)] *Biological and Soft Systems, Cavendish Laboratory, University of Cambridge, JJ Thomson Avenue, Cambridge, CB3 0HE, United Kingdom*

[2)] *Interdisciplinary Center for Nonlinear Phenomena and Complex Systems & Service de Physique des Systèmes Complexes et Mécanique Statistique, Université libre de Bruxelles (ULB), Campus Plaine, CP 231, Blvd du Triomphe, B-1050 Brussels, Belgium.*

[a)] Electronic mail: bmognett@ulb.ac.be

S1. REACTIONS BETWEEN BINDERS OF THE SAME TYPE

In this section we relax the hypothesis by which each complex cannot feature more than a single linker of a given type and re–derive Eq. 7 of the main text. When specifying a given complex made by i_1, i_2, $\cdots$, i_m linkers ($X = \{i_1, \cdots i_m\}$), we now only assume that $i_1 \leq i_2 \cdots \leq i_m$. The number of linkers of type i entering the complex $X = \{i_1, \cdots, i_m\}$ is defined as

$$g_i(X) = g_i(\{i_1, \cdots i_m\}) = \sum_{\alpha=1}^{m} \delta_{i,i_\alpha} \ . \tag{S1}$$

In the following with $\{\{X\}\}$ we refer to the ensemble of all possible complexes with at least two linkers while with $\{X\}_m$ we refer to the ensemble of complexes made by m linkers. Using these definitions it is not difficult to show that the partition function of the system (Eq. 5, main text) becomes

$$Z = \sum_{\{\{n_X\}\}} \prod_{i=1}^{c} \frac{N_i!}{n_i!} \prod_{m\geq 2} \prod_{\{X\}_m} \frac{1}{n_X! \left[\prod_{j=1}^{c} g_j(X)!\right]^{n_X}} \exp[-n_X \beta \Delta G_X] \ . \tag{S2}$$

Note that in Eq. S2 we do not distinguish between the g_i monomers in a given complex. This is not justified in systems featuring structured complexes where identical monomers can bind with different free energy depending on the site they occupy within the complex. This scenario may occur in nucleic acid complexes featuring several strands[1], isomeric clusters in gelation theory[2], or polymerization[3]. For the purpose of the present work this scenario would not change the final result in view of the fact that we derive an expression for the free energy of the system in which the binding free energy of the single complex ΔG_X is expressed in terms of equilibrium densities (see Eq. S5) and of the fact that different combinatorial factors of the complexes would simply re-define ΔG_X in Eq. S2.

Using Eq. S2 we can calculate the functional $\mathcal{A}$ defined in Eq. 5 of the main text

$$\beta\mathcal{A}(\{\{n_X\}\}) = \sum_{i=1}^{c} [n_i \log n_i - n_i - N_i \log N_i + N_i] + \sum_{m\geq 2} \left[\sum_{\{X\}_m} \Big[n_X \log n_X - n_X \right.$$
$$\left. + n_X \log \Big(\prod_{j=1}^{c} g_j(X)! \Big) + n_X \beta \Delta G_X \Big] \right] \tag{S3}$$

where the number of free binders of type i (n_i) is written as

$$n_i = N_i - \sum_{\{X\}} g_i(X) n_X \, . \tag{S4}$$

The stationary equations $\delta\mathcal{A}/\delta n_X = 0$ providing the equilibrium distribution $\overline{n}_X$ are $(\forall\, X)$

$$\beta\Delta G_X + \log\Big(\prod_{i=1}^{c} g_i(X)!\Big) + \log\overline{n}_X - \sum_{i=1}^{c} g_i(X)\log\overline{n}_i = 0\,. \tag{S5}$$

where we used Eq. S4. Notice in particular that Eqs. S5 can be rewritten into a standard equilibrium balance

$$\frac{\overline{n}_X}{\prod_{i=1}^{c}\overline{n}_i^{g_i(X)}} = \frac{e^{-\beta\Delta G_X}}{\prod_{i=1}^{c} g_i(X)!}\,. \tag{S6}$$

Using Eq. S5 multiplied by $\overline{n}_X$ in Eq. S3 we obtain the free energy of the system as a function of equilibrium distributions $\overline{n}_X$

$$\begin{aligned}
\beta F &= \sum_{i=1}^{c}\left[\overline{n}_i\log\overline{n}_i - \overline{n}_i - N_i\log N_i + N_i\right] + \sum_{m\geq 2}\left[\sum_{\{X\}_m}\left[\overline{n}_X\sum_{i=1}^{c} g_i(X)\log\overline{n}_i - \overline{n}_X\right]\right] \\
&= \sum_{i=1}^{c}\left[\left[\overline{n}_i + \sum_{\{X\}}\overline{n}_X g_i(X)\right]\log\overline{n}_i - N_i\log N_i + [N_i - \overline{n}_i]\right] - \sum_{\{\{X\}\}}\overline{n}_X \\
&= \sum_{i=1}^{c} N_i\log\frac{\overline{n}_i}{N_i} + \sum_{i}\sum_{\{X\}} g_i(X)\overline{n}_X - \sum_{\{\{X\}\}}\overline{n}_X \\
&= \sum_{i=1}^{c} N_i\log\frac{\overline{n}_i}{N_i} + \sum_{m\geq 2}\sum_{\{\{X\}\}_m}(m-1)\overline{n}_X
\end{aligned} \tag{S7}$$

where we have used multiple times Eq. S4 and the fact that $\sum_i g_i(X) = m$ if $X \in \{X\}_m$. Note that Eq. S7 has the same functional form of Eq. 7 of the main text.

S2. BINDING COOPERATIVITY IN DNA-FUNCTIONALIZED PARTICLES

We define by $\alpha(R_1, R_2, d)$ the volume of the intersection between two spheres of radius R_1 and R_2 with their center placed at distance equal to d. Defining $\sigma_+ = R_1 + R_2$ and $\sigma_- = |R_2 - R_1|$ we have

$$\alpha(R_1, R_2, d) = \frac{\pi}{12d}(\sigma_+ - d)^2(d^2 + 2d\sigma_+ - 3\sigma_-^2) \qquad\qquad \sigma_- < d < \sigma_+ \tag{S8}$$

Using the previous equation we can calculate the repulsive part of the potential due to entropic compression of the tethered DNA linkers[4]. In particular we find

$$\beta F_{\rm rep} = -3N\log\left[1 - n\frac{\alpha(R+L, R, d)}{\Omega_\infty}\right] - Nn\log\left[1 - \frac{\alpha(R+L, R, d)}{\Omega_\infty}\right] \tag{S9}$$

In Eqs. (9-11) of the main text v_A, v_B, and v_{AB} are the volume available to the sticky ends free to move on particle A, on particles B_i (when close to particle A), and when bridging particle A with particles B_i respectively. In particular we find

$$\begin{aligned} v_A &= \Omega_\infty - n\alpha(R+L,R,d) \\ v_B &= \Omega_\infty - \alpha(R+L,R,d) \\ v_{AB} &= \alpha(R+L,R+L,d) - 2\alpha(R+L,R,d) \end{aligned} \tag{S10}$$

where, as defined in the main text, n is the number of particles B_i attached to particle A, and Ω_∞ is the space availbale to the sticky ends on isolated particles

$$\Omega_\infty = \frac{4\pi}{3}\left[(R+L)^3 - R^3\right] . \tag{S11}$$

Note that Eqs. S9, S10, and S11 have been derived in the limit of $L \ll R$ and for double stranded DNA spacers modelled as rigid rods. Only when these assumptions hold the sticky ends are uniformly distributed in the layer between two spheres or radii R and $R+\ell$ [4]. For further geometrical assumptions we refer to the SI of Ref.[4]. If we define

$$\Xi_\ell(d) = \frac{e^{-\Delta G^0_\ell}}{\rho_\ominus v_A} \qquad \Xi_s(d) = \frac{e^{-\Delta G^0_s}}{(\rho_\ominus v_A)^2} \qquad \Xi_b(d) = \frac{v_{AB} e^{-\Delta G^0_b}}{\rho_\ominus v_A v_B} \tag{S12}$$

Eqs. (9–13) of the main text can then be rewritten as (assuming $i \neq j$, $i \neq k$, and $j \neq k$)

$$\begin{aligned} \overline{n}_{A_i} &= \frac{N}{1 + (\overline{n}_{A_j} + \overline{n}_{A_k})\Xi_\ell + \overline{n}_{A_j}\overline{n}_{A_k}\Xi_s + \epsilon_i \overline{n}_{B_i}\Xi_b} \\ \overline{n}_{B_i} &= \frac{N}{1 + \epsilon_i \overline{n}_{A_i}\Xi_b} \end{aligned} \tag{S13}$$

We numerically solve Eqs. S13 and use Eqs. 9-11 of the main text to calculate the fraction of hybridized strands. Results are given in Fig. S1 and Fig. 2**b** of the main text. In particular Fig. S1 reports the number of loops and spiders for an isolated A particle ($n = 0$) as given in Sec. II of the main text.

Appying Eq. 7 of the main text to this system we can then calculate the selective part of the interaction free energy

$$\beta F = N \sum_{i=1}^{3} \left[\log \frac{\overline{n}_{A_i}}{N} + \log \frac{\overline{n}_{B_i}}{N} \right] + \sum_{i=1}^{3} \left[\overline{n}_{\ell_i} + \overline{n}_{b_i} \right] + 2\overline{n}_s. \tag{S14}$$

The overall interaction free energy is then calculated by adding up Eq. S14 to the steric repulsion described by Eq. S9. The results are shown in Fig. 2**c** of the main text and Fig. S2

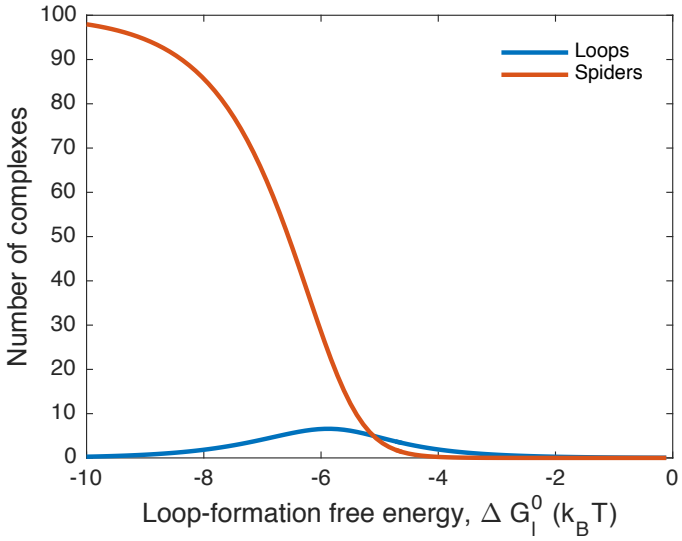

FIG. S1. Total number of loop and spider complexes as a function of ΔG^0_ℓ for an isolated A particle in which bridges cannot form. The free energy of the spider sticky-end complex is taken equal to $\Delta G^0_s = 3\Delta G^0_\ell$ as justified by the spider architecture suggested by Halverson and Tkachenko[5] formed by the hybridization of three complementary fragments of DNA, while a single hybridization directs the formation of loops. Note that such estimate neglects stacking terms and inert-tail effects that may be considerable.[6] Note also that it is easy to foresee more complex sticky-end designs that would allow to tune ΔG^0_ℓ and ΔG^0_s more independently.

S3. INTERACTION FREE ENERGY IN THE PRESENCE OF TOEHOLD-EXCHANGE-MECHANISM

The toehold system introduced in Ref.[7] and summarized in Sec. II of the main text features four types of two-strand complexes (see also Fig. 3 of Ref.[7]): $\ell_{1/2}$ are loops due to the hybridization of $\delta_{1/2}$ with γ, while $b_{1/2}$ are bridges due to the hybridization of $\delta_{1/2}$ with γ. The average number of two strand complexes is then given by

$$\begin{aligned}
\overline{n}_{\ell_1} &= \frac{\overline{n}_{\delta_1}\overline{n}_\gamma}{\rho_\ominus LA}\exp[-\beta\Delta G^0_{\gamma\delta_1}] \\
\overline{n}_{\ell_2} &= \frac{\overline{n}_{\delta_2}\overline{n}_\gamma}{\rho_\ominus LA}\exp[-\beta\Delta G^0_{\gamma\delta_2}] \\
\overline{n}_{b_1} &= \frac{\overline{n}_{\delta_1}\overline{n}_\gamma}{\rho_\ominus LA}\left(2-\frac{h}{L}\right)\exp[-\beta\Delta G^0_{\gamma\delta_1}] \\
\overline{n}_{b_2} &= \frac{\overline{n}_{\delta_2}\overline{n}_\gamma}{\rho_\ominus LA}\left(2-\frac{h}{L}\right)\exp[-\beta\Delta G^0_{\gamma\delta_2}]
\end{aligned} \tag{S15}$$

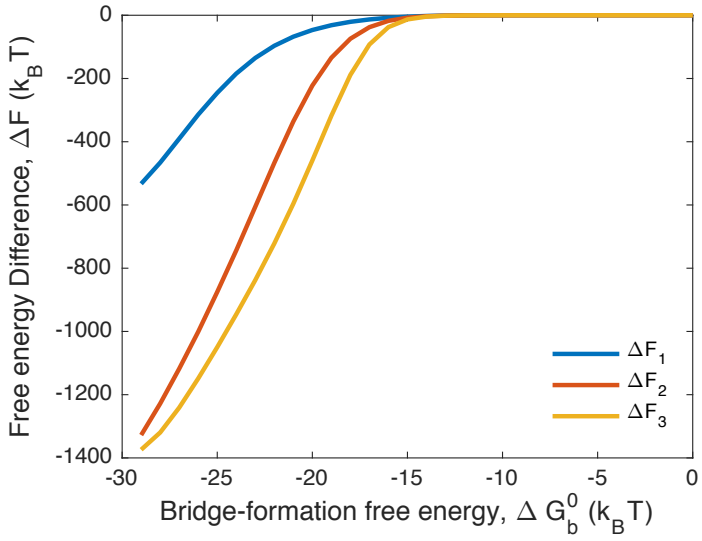

FIG. S2. Free energy difference between the complexes shown in Fig. 2**a** of the main text with $d = 2R + L$ as a function of the hybridization free energy of the sticky ends responsible for the formation of bridges.

where ΔG^0 are the hybridization free energies of the free sticky-ends in solution (we refer to the SI of Ref.[7] for their value). Following Ref.[7] we take the inter-membrane distance as $h = 1.4L$. The bottom panel of Fig. 3**d** in the main text reports the amount of two-strand complexes *per* γ strand

$$f_2 = \frac{\overline{n}_{b_1} + \overline{n}_{b_2} + \overline{n}_{\ell_1} + \overline{n}_{\ell_2}}{N} \tag{S16}$$

For isolated vesicles the bridge complexes are not possible and we have

$$\begin{aligned} \overline{n}^0_{\ell_1} &= \frac{\overline{n}^0_{A_1}\overline{n}^0_B}{\rho_\ominus LA} \exp[-\beta\Delta G^0_{\gamma\delta_1}] \\ \overline{n}^0_{\ell_2} &= \frac{\overline{n}^0_{A_2}\overline{n}^0_B}{\rho_\ominus LA} \exp[-\beta\Delta G^0_{\gamma\delta_2}] \end{aligned} \tag{S17}$$

We have four types of three-strand complexes, three bridging the two vesicles (t_1, t_2, and t_B) and the fourth (t_3) featuring a double loop structure (see Fig. 3 of Ref.[7]). The average number of complexes is then given by

$$\begin{aligned} \overline{n}_{t_1} = \overline{n}_{t_2} = \overline{n}_{t_B} &= m\frac{\overline{n}_{A_1}\overline{n}_{A_2}\overline{n}_B}{(\rho_\ominus LA)^2}\left(2 - \frac{h}{L}\right)\exp[-\beta\Delta G^0_{\gamma\delta_1\delta_2}] \\ \overline{n}_{t_3} &= m\frac{\overline{n}_{A_1}\overline{n}_{A_2}\overline{n}_B}{(\rho_\ominus LA)^2}\exp[-\beta\Delta G^0_{\gamma\delta_1\delta_2}] \end{aligned} \tag{S18}$$

where m is a multiplicity factor that counts the number of iso-energetic states ($m = 5$ in our case[7]). The top panel of Fig. 3**d** in the main text reports the amount of three strand complexes *per* γ strand

$$f_3 = \frac{\overline{n}_{t_1} + \overline{n}_{t_2} + \overline{n}_{t_B} + \overline{n}_{t_3}}{N} \tag{S19}$$

For isolated vesicles only t_3 is present and we have

$$\overline{n}^0_{t_3} = m \frac{\overline{n}^0_{A_1} \overline{n}^0_{A_2} \overline{n}^0_B}{(\rho_\ominus LA)^2} \exp[-\beta \Delta G^0_{\gamma\delta_1\delta_2}] \,. \tag{S20}$$

By applying Eq. 7 of the main text we can finally calculate the free energy *per* γ strand of the system

$$\begin{aligned} \beta F_{\rm att} = \log \frac{\overline{n}_B}{N} + \chi_1 \log \frac{\overline{n}_{A_1}}{\chi_1 N} + \chi_2 \log \frac{\overline{n}_{A_2}}{\chi_2 N} + \frac{\overline{n}_{b_1}}{N} + \frac{\overline{n}_{b_2}}{N} + \frac{\overline{n}_{\ell_1}}{N} + \frac{\overline{n}_{\ell_2}}{N} \\ + 2\frac{\overline{n}_{t_1}}{N} + 2\frac{\overline{n}_{t_2}}{N} + 2\frac{\overline{n}_{t_B}}{N} + 2\frac{\overline{n}_{t_3}}{N} \end{aligned} \tag{S21}$$

On the other hand for isolated vesicles we have

$$\beta F^0_{\rm att}(h = \infty) = \log \frac{\overline{n}^0_B}{N} + \chi_1 \log \frac{\overline{n}^0_{A_1}}{\chi_1 N} + \chi_2 \log \frac{\overline{n}^0_{A_2}}{\chi_2 N} + \frac{\overline{n}^0_{\ell_1}}{N} + \frac{\overline{n}^0_{\ell_2}}{N} + 2\frac{\overline{n}^0_{t_3}}{N} \tag{S22}$$

Finally in Fig. 3**c** of the main text we report the interaction free energy given by $\beta F_{\rm att} - \beta F^0_{\rm att}$.

REFERENCES

[1] R. M. Dirks, J. S. Bois, J. M. Schaeffer, E. Winfree, and N. A. Pierce, "Thermodynamic analysis of interacting nucleic acid strands," Siam Rev. **49**, 65–88 (2007).

[2] M. Gordon and G. Scantlebury, "Non-random polycondensation: statistical theory of the substitution effect," Trans. Faraday Soc. **60**, 604–621 (1964).

[3] D. Andrieux and P. Gaspard, "Nonequilibrium generation of information in copolymerization processes," Proc. Natl. Acad. Sci. USA **105**, 9516–9521 (2008).

[4] S. Angioletti-Uberti, P. Varilly, B. M. Mognetti, and D. Frenkel, ""Mobile linkers on DNA-coated colloids: valency without patches"," Phys. Rev. Lett. **113**, 128303–128306 (2014).

[5] J. D. Halverson and A. V. Tkachenko, "Sequential programmable self-assembly: Role of cooperative interactions," J. Chem. Phys. **114**, 094903 (2016).

[6]L. Di Michele, B. M. Mognetti, T. Yanagishima, P. Varilly, Z. Ruff, D. Frenkel, and E. Eiser, "Effect of inert tails on the thermodynamics of dna hybridization," J. Am. Chem. Soc. **136**, 6538–6541 (2014).

[7]L. Parolini, J. Kotar, L. Di Michele, and B. M. Mognetti, "Controlling self-assembly kinetics of DNA-functionalized liposomes using toehold exchange mechanism," ACS Nano **10**, 2392–2398 (2016).