ACS Macro Letters

Letter

pubs.acs.org/macroletters

Network Formation by Cross-Hybridization of Complementary Strands to Grafted ssDNA

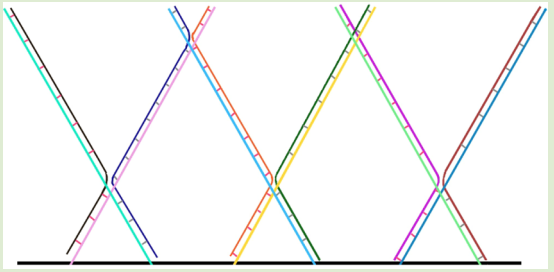
S. Medalion,[†] M. Wagman,[†] A. Y. Grosberg,[‡] and Y. Rabin*[,†]

[†]Department of Physics and Institute for Nanotechnology and Advanced Materials, Bar−Ilan University, Ramat-Gan 52900, Israel

[‡]Department of Physics and Center for Soft Matter Research, New York University, New York, New York 10003, United States

Supporting Information

ABSTRACT: When a low density brush of single-stranded DNA (ssDNA) targets end-grafted to a surface is immersed in a solution of complementary ssDNA probes, a regular brush of DNA duplexes is formed by 1:1 hybridization between probe and target DNA. We suggest that in higher density brushes of ssDNA this process competes with cross-hybridization of a target strand to several neighboring probe strands resulting in the formation of a cross-linked DNA network. We analyze a simple 2D model of a dense DNA brush and use analytic methods and computer simulations to find how the conditions for network formation depend on system size and DNA length. We argue that in 3D brushes cross-hybridization will nearly always lead to network formation and suggest that this may explain some intriguing results on dense DNA brushes. Experiments on DNA monolayers and concentrated DNA solutions that could test our predictions are proposed.



Hybridization of complementary (target) strands to surface-grafted (probe) single-stranded DNA (ssDNA) is used in all microarray-based methods for DNA analysis.[1−4] In these applications the density of the grafted molecules is sufficiently low so that each target strand hybridizes with a single probe strand.

In this communication we propose that in dense DNA brushes where the distance between the grafting points is much smaller than the radius of gyration of isolated strands a target strand can hybridize to two or more neighboring probe strands by switching from one to the other and forming an interstrand junction (see Figure 1) that resembles a Holliday junction in genetic recombination[5] and that under suitable conditions a macroscopic network of cross-hybridized DNA molecules can form. Since all grafted DNA molecules are identical, interstrand junctions can form anywhere along the DNA sequence (this degeneracy is lifted in DNA origami where sequences are predesigned to create a junction in a particular location on ssDNA[5,6]). Below we use a toy model of a DNA brush to study the conditions under which such a network is created.

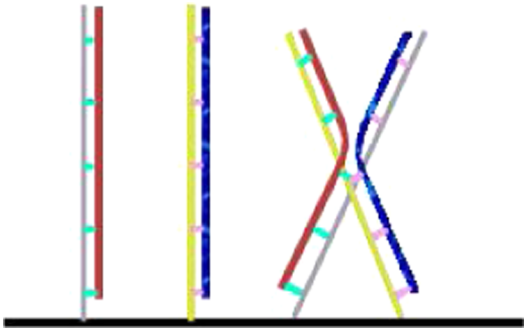


Figure 1. DNA brushes. Modes of hybridization in a high density brush: a target strand that hybridizes with a single probe strand (left and center) and two target strands each of which switches between two probe stands, thus connecting them and creating an interstrand junction (right).

To obtain analytic results we first consider a 2D polymer brush of $2k$ ssDNA molecules, grafted along a line. We assume that hybridization is complete (no unpaired DNA except at interstrand junctions) and that DNA oligomers are shorter than the Kuhn segment of double helix DNA (300 base pairs). Using mean-field arguments we show in the Supporting Information (SI) that grafted probe strands can be represented by straight lines of length l that lie on a 2D lattice which is completely covered by target strands (see Figure 2).

Inspection of Figure 2 shows that there are two kinds of nodes on the lattice: (1) interstrand junction (junction, for brevity) in which each of the two target strands that enter the node hybridizes with two probe strands and therefore changes direction at the crossing point (junctions are indicated by circles in Figure 2) and (2) pseudocrossing (crossing, for brevity) in which each of the target strands is hybridized to the same probe strand on both sides of the node and therefore maintains its direction at the crossing. Thus, a junction represents a connection between two chains, while a crossing represents two chains that are not connected at this point. All

Received: December 2, 2013
Accepted: January 24, 2014
Published: January 29, 2014

   dx.doi.org/10.1021/mz400622a | ACS Macro Lett. 2014, 3, 191−193

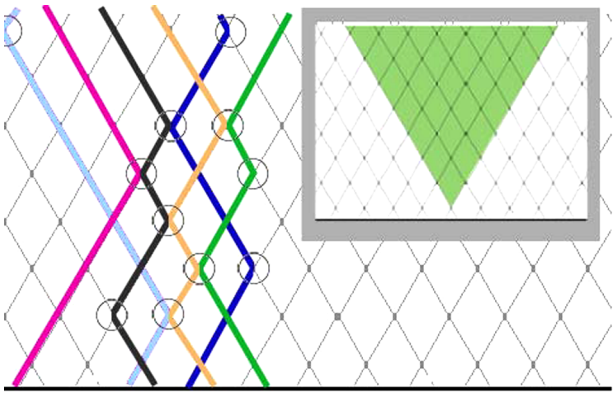


Figure 2. Hybridization on a 2D lattice. A lattice formed by grafted ssDNA probes of $n = 8$ segments (thin gray lines), with 6 complementary ssDNA targets hybridizing to them in various conformations (thick colored lines). Circles represent interstrand junctions. Other hybridized target strands are not shown. Inset: the green inverted triangle represents a region completely filled with pseudocrossings that generates a break in the network.

possible states of hybridization are uniquely determined by specifying the state of each of the nk lattice nodes: a crossing or a junction. The DNA sequence is assumed to be sufficiently nondegenerate to ensure that it is a unique function of the distance from the surface (z), so that target DNA that begins to hybridize with a probe starting from the bottom of the lattice and then switches to another probe strand can only go up and therefore behaves as a directed random walk. In this nondegenerate case even a single step of length a in the $-z$ direction would lead to mismatch between the bases of the target and the probe, and taking the distance between crossing points as greater than or equal to a DNA period of 10.5 bp (to ensure stable interstrand hybridization), such a mismatch would carry a prohibitively large energy penalty. Assuming that each molecule is made of n segments of length a ($l = na$), a hybridized target strand will be represented either by a straight line (regular DNA duplex) or by a broken one (interstrand hybridization); in both cases its two ends will be located at the bottom and the top of the brush, respectively (Figure 2).

Since the formation of an interstrand junction carries energy penalty ε due to bending, stacking, and base pairing defects, the probability of creating such a junction is $p = \exp(-\varepsilon/k_BT)/[1 + \exp(-\varepsilon/k_BT)] < 1/2$. When p is sufficiently small, interstrand hybridization is negligible, and most DNA will form duplexes. Upon increasing p, the number of "cross-links" between adjacent probe strands will increase until the gelation threshold is reached at $p = p_{gel}$, and a connected network that spans the entire system (such that in a typical realization there is at least one cross-link between any of $2k$ grafted chains) will form. Note that while our problem resembles percolation on a 2D lattice,[7] it differs from it in that each hybridized probe can contain up to n junctions, and therefore, to introduce a break in the network one would have to generate a region of thickness n that does not contain any junctions. A more careful analysis shows that the 2D network will break into two disconnected parts when a triangle or an inverted triangle (the latter is shown in the inset to Figure 2) that does not contain any junctions appears in the lattice. The value of p_{gel} can be estimated as follows: if one starts with a fully connected network and decreases p, the number of junctions will decrease until a triangle that does not contain any junctions appears and the network is broken. Since such a triangle contains $n(n + 1)/2$ crossings and its apex can be at any of k locations, for k not exponentially large compared to n the probability of its formation is $2k(1 - p)^{n(n+1)/2}$ (the factor of 2 reflects the fact that the triangle can be normal or inverted). At $p = p_{gel}$ this probability approaches unity, and we obtain an expression for p_{gel} in terms of length and number of grafted molecules

$$p_{gel} = 1 - (2k)^{-2/[n(n+1)]} \quad (1)$$

Inspection of this expression suggests that as system size increases p_{gel} becomes larger than 1/2 and a network that spans the entire system cannot form. This concurs with the expectation that as $k/n \to \infty$ our problem is reduced to percolation in 1D for which $p_{gel} = 1$. However, since p_{gel} decreases even more rapidly with DNA length, this is not necessarily the case. Indeed for $n^2 \gg \ln k$, eq 1 can be approximated as $p_{gel} \simeq (2/n(n + 1))\ln(2k) \ll 1$. For example, taking the spacing between grafting points to be 4 nm (grafting density[8] 6×10^{12} DNA/cm^2) and "surface" dimension of 20 μm gives $k = 5000$. For $n = 30$ we get $p_{gel} = 0.02$, and therefore a 20 μm long network will form if the free energy cost of an interstrand junction is $\varepsilon \leq 3.9\ k_BT$. Previous estimates yield about $2k_BT$ per junction,[5] and we conclude that the conditions for network formation can be satisfied in our toy model.

Next, we performed computer simulations in which for each pair of values (n,k) a lattice was generated and its nodes were randomly labeled as crossings (with probability $1 - p$) or as junctions (with probability p). The process was repeated 10 000 times, and in each realization we checked for the appearance of a break (triangle without junctions). We then calculated $p = p_{gel}$ for which there was one triangle on the average in our ensemble of realizations. Comparison between eq 1 and results of computer simulations is shown in Figure 3.

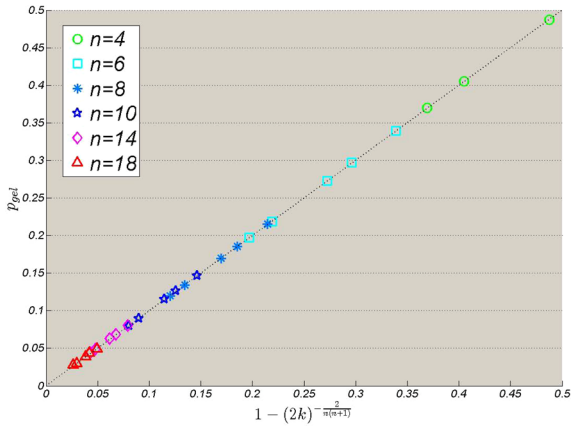


Figure 3. Network formation threshold. Comparison between eq 1 (solid line) and simulation results for p_{gel} as a function of system size k (in the range 50−5000) and DNA length n (in the range 4−18).

So far, we did not carry out computer simulations of network formation in a 3D brush; nevertheless, we can estimate p_{gel} in the spirit of the arguments that led to eq 1. To generate a break in a 3D network one has to separate the two sides of the $k \times k$ surface by a wall that does not contain any junctions, of thickness and height n, and of some length L (a local defect such as the triangle shown in Figure 2 does not affect the global connectivity of the 3D network). Such a wall can be thought of as a self-avoiding walk of length L and step n, and the

dx.doi.org/10.1021/mz400622a | ACS Macro Lett. 2014, 3, 191−193

probability of its realization is proportional to $z^L(1-p)^{n^2L}$ where z^L represents the number of walks of length L (up to power law corrections to the main exponential). Since we have replaced the wall by a line, the effective coordination number z is of the order z_0n^2 (z_0 is the coordination number of the lattice of grafting points), and the probability can be written as $\exp\{L\cdot \ln[z_0n^2(1-p)^{n^2}]\}$. This probability decreases exponentially with L (the argument of the logarithm is smaller than unity), and therefore the sum over the paths is dominated by the shortest ones—straight lines of length $L \simeq k$. Such a wall will contain n^2k crossings, and since $n^2k \gg \ln k$ always holds, this yields

$$p_{\mathrm{gel}}^{3\mathrm{D}} \simeq \frac{\ln k}{n^2 k} \ll 1 \tag{2}$$

and we conclude that cross hybridization in the 3D brush will lead to formation of a network even when the probability of forming an interstrand junction is very small.

We presented a theoretical analysis of network formation by spontaneous cross-hybridization in the DNA brush. The mechanism of network formation is purely entropic: for sufficiently long DNA chains, the entropy associated with a large number of ways of forming a network can overcome the free energy cost of cross-hybridization. While our exact results were obtained for a 2D lattice model, the prediction that probability of gelation (for fixed p) increases with DNA length is physically intuitive and is expected to be robust. We presented a tentative estimate which suggests that such networks will form even under less restrictive conditions in dense 3D DNA brushes. We plan to test these results by off-lattice simulations using coarse-grained models of DNA hybridization.[9]

To test our predictions experimentally, one would have to carry out nanomechanical studies[10] that can distinguish between regular and cross-linked DNA monolayers: the shear modulus of a grafted network is expected to be larger than that of a regular brush; the height of a DNA network should be smaller that that of a brush of DNA duplexes. Notice that in addition to DNA length and grafting density, experiments may also be sensitive to temperature and sequence since the free energy of cross-hybridization depends on both. As there is only a single lowest energy state that corresponds to regular hybridization and formation of DNA duplexes and a huge number of higher energy states that correspond to different network structures formed by interstrand hybridization, the latter may dominate the kinetics of hybridization in experiments on dense ssDNA brushes. Indeed, kinetic effects resulting from hindered diffusion of target DNA into the forming network can explain the observed lowering of hybridization efficiency with increasing ssDNA monolayer density.[8] Network formation may also explain the change from intermolecular repulsion to attraction (despite the higher Coulomb repulsion) following hybridization of grafted probe ssDNA with complementary target ssDNA oligomers, observed by nanomechanical cantilever motion.[11] Another intriguing possibility is that networks may form in dense dsDNA solutions that are denatured by heating above the melting temperature of DNA (denaturing the duplexes) and then annealed by cooling (forming a network). Such DNA networks are expected to be metastable upon dilution in water and can be directly observed by fluorescence from intercalating dyes and by other methods.

■ ASSOCIATED CONTENT

Supporting Information

Mean-field arguments leading to the 2D lattice model (Figure 2) are given. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: yitzhak.rabin@biu.ac.il.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by grants from the US-Israel Binational Science Foundation (YR and AYG) and by the I-CORE Program of the Planning and Budgeting committee and the Israel Science Foundation (YR).

■ REFERENCES

(1) Maskos, U.; Southern, E. M. *Nucleic Acids Res.* **1993**, *218*, 4663−4669.

(2) Lockhart, D. J.; Dong, H. L.; Byrne, M. C.; Follettie, M. T.; Gallo, M. V.; Chee, M. S.; Mittmann, M.; Wang, C. W.; Kobayashi, M.; Horton, H.; Brown, E. L. *Nat. Biotechnol.* **1996**, *14*, 1675−1680.

(3) Chan, V.; Graves, C. J.; McKenzie, S. E. *Biophys. J.* **1995**, *69*, 2243−2255.

(4) Wang, J. *Nucleic Acids Res.* **2000**, *28*, 3011−3016.

(5) Lu, M.; Marky, L. A.; Seeman, N. C.; Kallenbach, N. R. *J. Mol. Biol.* **1992**, *223*, 781−789.

(6) Rothemund, P. W. K. *Nature* **2006**, *440*, 297−302.

(7) Stauffer, D.; Aharony, A. *Introduction to Percolation Theory*; Taylor and Francis: UK, 1994.

(8) Peterson, A. W.; Heaton, R. J.; Georgiadis, R. M. *Nucleic Acids Res.* **2001**, *29*, 5163−5168.

(9) Hinckley, D. M.; Freeman, G. S.; Whitmer, J. K.; de Pablo, J. J. *J. Chem. Phys.* **2013**, *139*, 144903.

(10) Kosaka, P. M.; Gonzalez, S.; Dominiquez, C. M.; Cebollada, A.; San Paulo, A.; Calleja, M.; Tamayo, J. *Nanoscale* **2013**, *5*, 7425−7432.

(11) Wu, G.; Haifeng, J.; Hansen, K.; Thundat, T.; Datar, R.; Cote, R.; Hagan, M. F.; Chakraborty, A. K.; Majumdar, A. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *98*, 1560−1564.