Electronic Supplementary Material (ESI) for Nanoscale
This journal is © The Royal Society of Chemistry 2013

Supporting Information

Structural and Energetic Basis for Hybridization Limits in High-Density DNA Monolayers

*Giovanni Doni,[a] Maryse D. Nkoua Ngavouka,[b,c] Alessandro Barducci,[d] Pietro Parisse,[c] Alessandro De Vita,[a] Giacinto Scoles,[c,e] Loredana Casalis,[c] and Giovanni M. Pavan[f,]**

a: Department of Physics, King's College London, London WC2R 2LS, UK

b: University of Trieste, PhD School in Nanotechnology and Nanoscience, Piazzale Europa,1 34127 Trieste, IT

c: Elettra-Sincrotrone Trieste S.C.p.A, INSTM-ST unit, Strada Statale 14—km 163,5 in AREA Science Park, I-34149 Basovizza, Trieste, IT

d: Laboratory of Statistical Biophysics, ITP SB EPFL, CH-1015 Lausanne, CH

e: Department of Biological and Medical Science, University of Udine, and Azienda Ospedaliera Universitaria Piazza Santa Maria della Misericordia, 33100 Udine, IT

f: Department of Innovative Technologies, University of Applied Science of Southern Switzerland, Galleria 2, 6928 Manno, CH

giovanni.pavan@supsi.ch

TABLE OF CONTENTS

This journal is © The Royal Society of Chemistry 2013

Experimental protocol

Chemicals. The thiolated sequence (SH-$(CH_2)_6$-5'-TAATCGGCTCATACTCTGACTG-3') and the complementary hybrid sequence (5'-CAGTCAGAGTATGAGCCGATTA-3') were purchased from Biomers GmbH (HPLC purification grade). Top oligo ethylene-glycol-terminated alkylthiols (HS-$(CH_2)_{11}$-$(OCH_2CH_2)_6$-OH: TOEG6) were purchased from Sigma Aldrich.

AFM. Atomic Force Microscope (AFM) measurements have been carried out on a XE-100 PARK Instruments equipment. Nanografting and imaging experiments have been performed in contact mode on a custom liquid cell with standard silicon cantilevers (Micromasch NSC19 and NSC38, spring constant 0.6 and 0.03 N/m, tip radius curvature < 10 nm). For imaging experiments, the topographic signals were recorded in TE buffer (200 mM NaCl and 400 mM NaCl) at low force (~0.1 nN) in order to avoid compression of the DNA layer.

Preparation of TOEG6 SAM and nanografting of ssDNA. Freshly cleaved ultraflat gold samples were soaked into a freshly prepared 300 μM solution of TOEG6 in ethanol and kept in dark at 4 C for 12-18 hours. The resulting SAM was then rinsed with ethanol, dried under a soft stream of nitrogen and immediately glued into the AFM liquid cell. According to the protocol of nanografting reported earlier,[S1-S5] we scan at high load (80-100nN) an AFM tip on a selected area (micrometer sized) in the presence of thiolated ss-DNA (5μM) dissolved in a TE 1M NaCl solution. The high load applied is locally promoting the replacement of the TOEG6 molecules from the surface with thiolated DNA molecules present in the solution. After resetting the force to lower values (~0.1 nN), the DNA nanostructure can be imaged, and the height of the nanografted patch referred to the SAM carpet can be accurately measured.

Hybridization reactions were carried out for 1 hour at room temperature within the AFM liquid cell filled with a TE 1M NaCl solution containing 1 μM complementary DNA ssDNA.

This journal is © The Royal Society of Chemistry 2013

Preparation of ssDNA SAM and nanografting of TOEG6. Freshly cleaved gold substrates were immersed in a 1M NaCl TE buffer solution with 1 µM thiolated ssDNA for 24 hours and subsequently immersed in a 1M NaCl TE buffer solution containing 1mM mercaptohexanol solution for 1 hour. TOEG6 has been nanografted into the freshly prepared ssDNA SAMs using a 3:2 mixture of 1M NaCl TE buffer and ethanol with 15µM of TOEG6 molecules.

Hybridization reactions were carried out for 1 hour at room temperature within the AFM liquid cell filled with a TE 1M NaCl solution containing 1 µM complementary DNA ssDNA.

In Figure S1 we report an example of our measurements on ssDNA monolayers before and after hybridization.

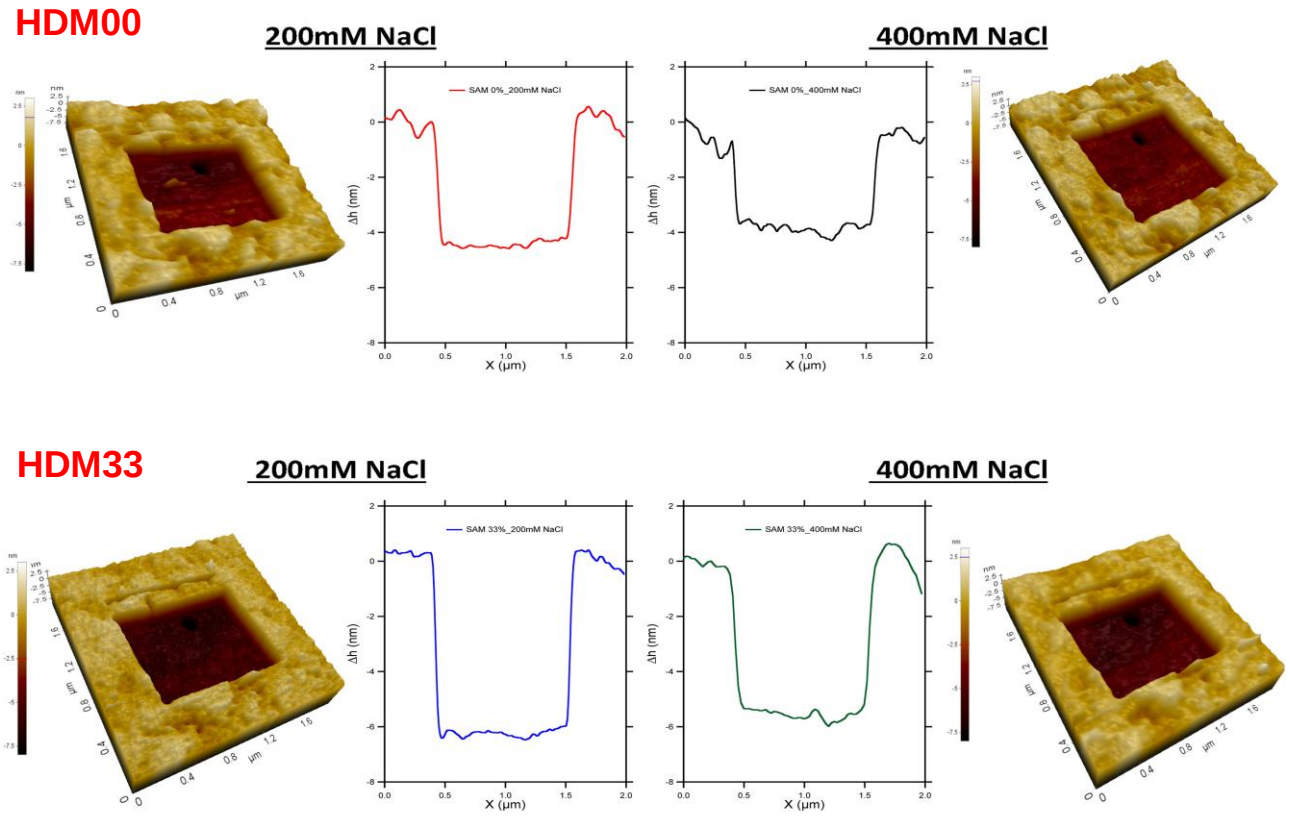

Figure S1. AFM micrographs of TOEG6 patches nanografted into a high density ssDNA dna monoloayers before (**HDM00**) and after (**HDM33**) hybridization imaged in different salt conditions (200mM NaCl and 400mM NaCl) and corresponding average height profiles.

This journal is © The Royal Society of Chemistry 2013

The grafting of a TOEG6 patch into a ssDNA **HDM** serves as a reference for the height of the DNA monolayer. We can clearly observe the variation of the relative height (Δh) changing the salt concentration in the imaging buffer accordingly with the simulations reported in the main text. Moreover also the variation due to the hybridization matches with the simulation for the used salt concentrations. The absolute height values reported in the main text are the values obtained averaging at least 3 different experiments.

Computational details

Creation of the molecular systems. The Au (111) surface was built as composed by five monoatomic Au layers, and served as a rhombic base for the simulation cell. A total of 3 x 3 thiolated dsDNA, composed of a thiolated ssDNA (sequence: S-$(CH_2)_6$-5'-TAATCGGCTCATACTCTGACTG-3') and the complementary hybrid (sequence: 5'-CAGTCAGAGTATGAGCCGATTA-3'), were grafted on the gold surface. The Sulfur atom of each of the 9 thiolated dsDNAs was bound to the first gold layer – *i.e.*, to single Au atoms separated each other by a distance of 3.1 nm. This DNA axial distance was consistent with the experimental density of 1.2 x 10^{13} molecules cm^{-2}. A periodic box with initial size of 97 x 97 x 183 Å (angles: 90° 90° 60°) was built around the solute (0 buffer on the *xy* plane – the simulation box base coincides with the rhombic gold surface). The simulation cell containing 9 thiolated dsDNA molecules and the rhombic Au (111) surface was filled with explicit TIP3P[S6] water molecules and Cl^- counterions to guarantee the system neutrality using the *leap* module of the AMBER 11 suite of programs.[S7] Water molecules were replaced with ions if eventual superposition occurred. The first HDM system was thus obtained – **HDM100**, composed by 9 hybridized dsDNA. The equilibrated box size of 97 x 97 x 150 Å (angles: 90° 90° 60°) led to a solution ionic strength of ≈300 mM NaCl, which was maintained as a constant for all the differently hybridized simulated systems.

This journal is © The Royal Society of Chemistry 2013

Table S1. The main features of the different HDM molecular systems simulated in this work.

HDM system	Number of dsDNA in the patch	Number of ssDNA in the patch	Total HDM charge[a] (e)	Number of Cl^- ions in the system	Number of Na^+ ions in the system	Corresponding ionic strength[b] (mM)
HDM00	0	9	-198	95	293	≈300
HDM11	1	8	-219	84	303	≈300
HDM22	2	7	-240	74	314	≈300
HDM33	3	6	-261	63	324	≈300
HDM44	4	5	-282	53	335	≈300
HDM56	5	4	-303	42	345	≈300
HDM67	6	3	-324	32	356	≈300
HDM78	7	2	-345	21	366	≈300
HDM89	8	1	-366	11	377	≈300
HDM100	9	0	-387	0	387	≈300
HDM00[c]	0	9	-198	32	230	≈200
HDM33[c]	3	6	-261	0	261	≈200
HDM00[d]	0	9	-198	159	357	≈400
HDM33[d]	3	6	-261	127	388	≈400

[a] Each 22-base pair grafted tiolated ssDNA carries a negative charge of -22 e. Each complementary strand carries a negative charge of -21 e since, by default, the last base of the strand is truncated and does not possess the P atom. Thus, each dsDNA carriesa negative charge of -43 e. [b] The minimum amount of Na^+ ions (387) necessary to neutralize the most negatively charged HDM system (**HDM100** – total charge -387 e) was initially added in solution, leading to a ionic concentration of ≈300 mM NaCl. [c,d] The number of ions in the system was changed to reproduce the experimental ionic strengths of ≈200 mM NaCl and ≈400 mM NaCl and to test the effect of salt concentration on the patch height before (**HDM00**) and after hybridization has occurred (**HDM33**).

The systems with lower hybridization percentage were obtained by deleting an increasing number of complementary hybrid sequences from the **HDM100** thiolated dsDNAs, consequently transformed into ssDNAs. The voids generated by the strands deletion were filled again with TIP3P water molecules using the *AddToBox* utility of AMBER 11. Since in these systems decreasing

This journal is © The Royal Society of Chemistry 2013

hybridization caused a decrease in the overall solute charge, some of the Cl^- ions were replaced by Na^+ in order to maintain constant the ionic strength in the box (≈300 mM NaCl) and, at the same time, to guarantee the system neutrality.

Once simulations have identified the **HDM33** system as representative of the HDM hybridized state, the reliability of the HDM model used in this study has been challenged further. **HDM33** and **HDM00** represented HDM before and after hybridization. A periodic box containing explicit water molecules and the suitable number of Cl^- and Na^+ ions necessary to reproduce the experimental ionic strengths of ≈200 mM and ≈400 mM NaCl was constructed according to the same procedure described previously. The number of Cl^- and Na^+ ions introduced in the ≈200 mM and ≈400 mM systems were calculated based on the simulation periodic box dimensions and volumes. Four additional systems – **HDM00** 200 mM, **HDM33** 200 mM, **HDM00** 400 mM and **HDM33** 400 mM were thus obtained and simulated as described in the next section.

Parametrization and simulation procedure: During all MD simulations the Au atoms composing the Au (111) surface were maintained as fixed. The force field parameters for the gold atoms and the alkanethiols were taken from literature.[S8] The DNA was parametrized according to the *parm99* all-atom force field by Cornell *et al.*,[S9] and accounting for the nucleic-acids force field improvements reported by the group of Orozco.[S10] The *parm99* all-atom force field was used for all the other standard residues within the systems.

All the MD simulations were conducted in periodic boundary conditions using AMBER 11.[S7] All systems were initially minimized. After this, a first steps of molecular dynamics simulation (MD) was run for 100 ps in NVT conditions to reach the simulation temperature of 300 K and to start relaxing the solvent inside the periodic box. During this phase, the whole solute was restrained and the solvent was let free to move. After this initial step, the density and the pressure of the HDM systems have been equilibrated for 8 ns of MD simulation at the temperature of 300 K, using a time step of 2 femtoseconds, the Langevin thermostat and a 8 Å cutoff. During this step, only the Au

atoms were maintained as fixed, and the system was pressurized (p = 1 atm) by adjusting the box size along the *z* axis. In this phase, the simulation periodic boxes reached the equilibrium dimensions. The particle mesh Ewald[S11] (PME) approach was adopted to treat the long-range electrostatic effects and the SHAKE algorithm was used on the bonds involving Hydrogen atoms[S12] for all the simulation steps. After these preliminary runs, each system was equilibrated by running a 100 ns NVT MD run during which the Au atoms were fixed. All HDM simulated systems reached the equilibrium with good stability during the first 50 ns of MD simulation (the root mean square deviation, rmsd, data obtained from the MD trajectories were monitored to verify the equilibrium – see Figure S2). All MD simulations were carried out using the *pmemd.cuda* module of AMBER 11 working on NVIDIA Tesla 2050 and GTX 580 GPU cards.

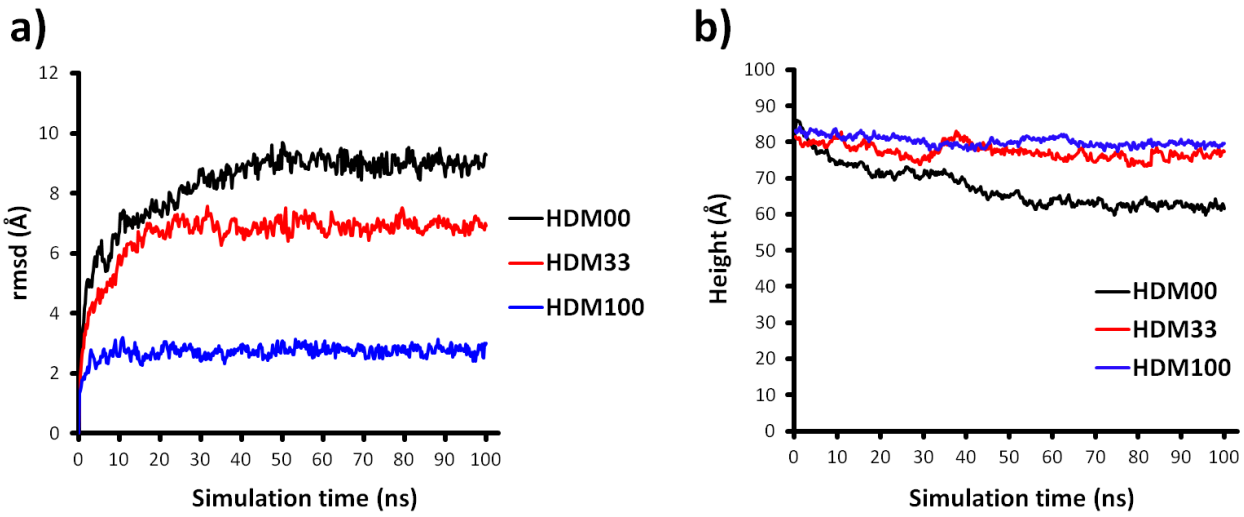

Figure S2. Structural data used to verify the systems equilibration during the MD simulations. (a) Root mean square data (rmsd) is expressed in Å and plotted as a function of simulation time (in ns). (b) HDM heights (in Å) are measured at each MD simulation time step as the average distance between the topmost base of each oligonucleotides within the patch and the Au surface. Data are reported for the **HDM00** (black), **HDM33** (red) and **HDM100** (blue).

Energetic analysis: The energy (E) for each system was calculated according to the MM-PBSA approach.[S13] Figure S3 shows the energy convergence during the MD simulations. For all

This journal is © The Royal Society of Chemistry 2013

simulated HDM systems, 250 snapshots taken from the equilibrated phase (the last 50 ns) of the MD trajectories were considered for the energetic analysis (Figure S3b). In the calculation of the energy (E) of the different solute systems, the gold surface was not considered, and only the energetic contributions of the DNA molecules were taken into account.

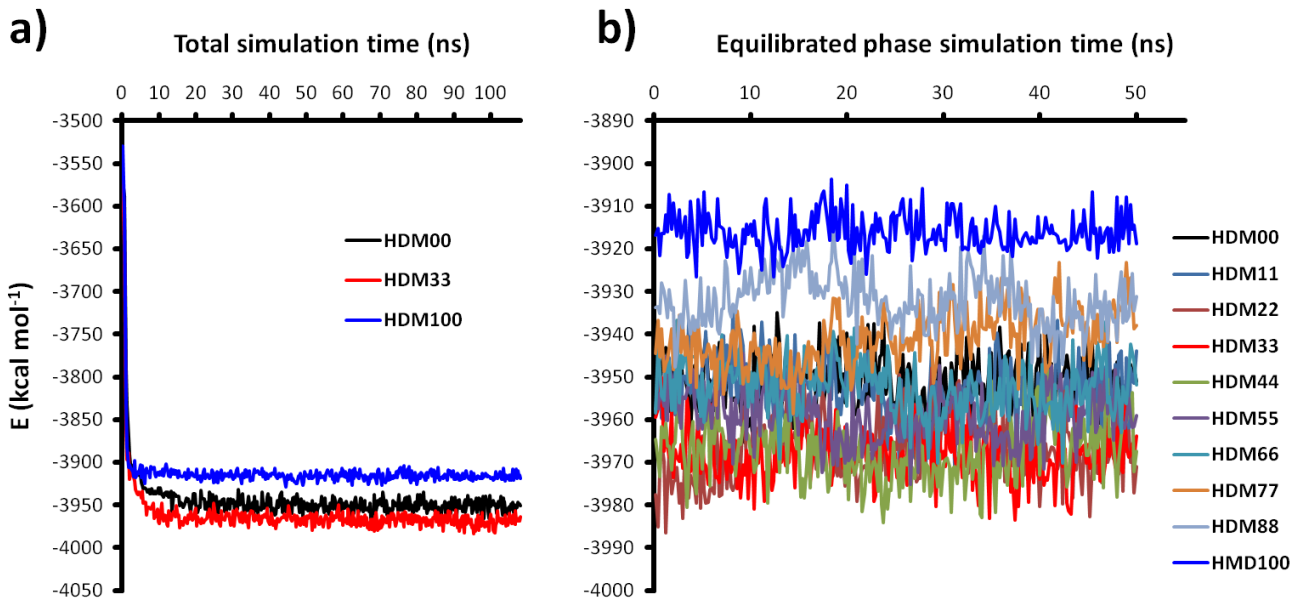

Figure S3. Energy (E) plots as a function of the simulation time obtained from the MD runs. (a) Convergence of the E profile during the entire MD simulation (full data are reported for **HDM00**, **HDM33** and **HDM99** systems). (b) The last 50 ns of the equilibrated phase MD simulations were used for the energetic analysis of all simulated HDM systems.

The interaction energy E is defined by Eq. (S1):

$$E = E_{gas} + E_{sol} \qquad \text{(S1)}$$

The energy (E) can be split into total gas-phase in vacuum non-bond energy (E_{gas}), composed by a coulombic and a van der Waals term (E_{ele} and E_{vdW}), and a solvation energetic term ($E_{sol} = E_{PB} + E_{NP}$)[S14] as described in Eq. (S2). The polar component of E_{PB} was evaluated using the Poisson-Boltzmann[S15] (PB) approach with a numerical solver implemented in the *pbsa* program of AMBER

11.[S16] The non-polar contribution to the solvation energy was calculated as $E_{NP} = \gamma$ (SASA) + β, in which $\gamma = 0.00542$ kcal/Å^2, $\beta = 0.92$ kcal/mol, and SASA is the solvent-accessible surface estimated with the MSMS program.[S17] Finally, the normal-mode[S18] approach was used to compute the entropic term. The calculated energies have been normalized per-DNA strand in order to allow comparison between differently hybridized HDM systems. The ΔE and $T\Delta S$ values obtained for each hybridized system were calculated respect to the energy of **HDM00** – system at 0% hybridization – that was here used as a reference (Figure S4a).

Further energetic analysis was performed by keeping the number of DNA strands into each HDM system constant (=18). Since the number of DNA strands changes in the patch along with differently hybridized systems **HDMnn**, a variable number of complementary unbound ssDNA strands were considered in the energetic analysis for each simulated system.

The **HDM100** system, for example, is composed of 9 grafted dsDNA, so that all 18 DNA strands fit into the patch. On the other hand, the **HDM33**, system is composed of 3 grafted dsNDA (6 strands), and 6 out of 9 total grafted chains are ssDNA for a total number of 12 DNA strands within the patch system. Thus, the contribution of 6 additional unbound complementary ssDNA free in solution was added for reaching the total of 18 DNA strands in the system in the energetic analysis.

A molecular system composed of a single complementary ssDNA strand immerged in a periodic box extending 14 Å from the ssDNA atoms, and containing explicit water molecules and the necessary number of Na^+ and Cl^- ions to reproduce the same salt concentration in solution used for the HDM simulated cases (≈300 mM NaCl), was equilibrated for 200 ns of NPT MD simulation at 300 K and 1 atm of pressure, according to the same protocol used for all the other simulations. Energy values for the additional free ssDNA strands in solution have been extracted from the MD equilibrated trajectories according to the same procedure adopted previously for the HDM systems.

Again, the total energies obtained for all HDM systems were normalized per-DNA strand for comparison, consistent with what was done in the previous energetic analysis. The energy of each final differently hybridized state were then compared with that of the initial reference non-

hybridized state **HDM00** (which, in this case, accounts for the presence of additional 9 unbound free ssDNA in solution) to obtain the energetic variations (ΔE and TΔS) associated to increasing hybridizations (Figure S4b).

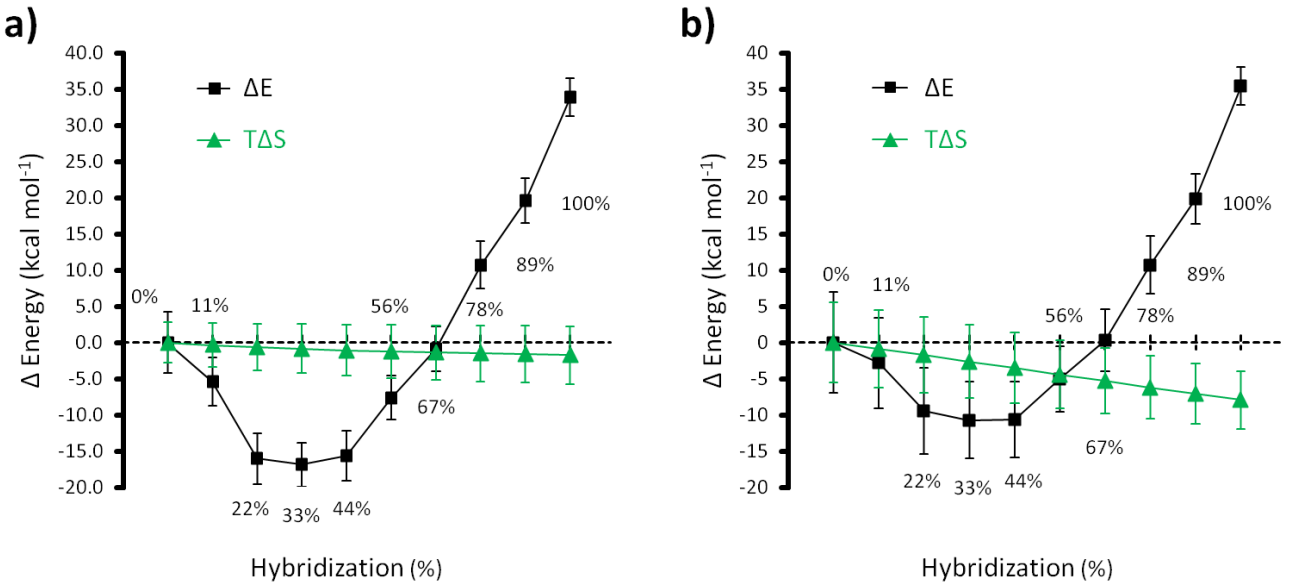

Figure S4. ΔE and TΔS values calculated for the differently hybridized **HDMnn** systems respect to the initial non-hybridized reference system **HDM00**. Energies were normalized per-ssDNA strand for comparison. ΔE and TΔS variations are expressed in kcal mol^{-1} (dark squares are related to ΔE and green triangles to TΔS). (a) Energy variations related to the **HDMnn** patch systems (the number of strands changes in the patch). (b) Energy variations related to the HDM systems considering the presence of a variable number free unbound complementary for maintaining constant the total number of DNA strands in the system (18 patch+unbound free complementary ssDNA strands).

This additional analysis demonstrates that the shape of the ΔE curve does not change substantially when accounting for the free complementary ssDNA in solution. Intermolecular repulsion plays a major role in controlling the HDM hybridization limit. TΔS decreases more strongly in Figure S4b than in Figure S4a due to the fact that free ssDNA have higher entropy that is lost upon hybridization.

This journal is © The Royal Society of Chemistry 2013

Structural analysis: The heights and the density profiles of **HDM00** and **HDM33** model systems at 200 mM and 400 mM NaCl were extracted from the equilibrated phase (last 50 ns) MD trajectories. The distance between the topmost atom of all DNA (double or single-stranded) and the upper Au(111) layer was calculated for each model. Figure S2 reports the individual DNA heights for the four systems extracted from the equilibrated phase of the MD trajectories. The patch heights were calculated as the average of height values of the single chains in the simulation box. This assumption found consistency with the height extracted from the density profiles (see further discussion).

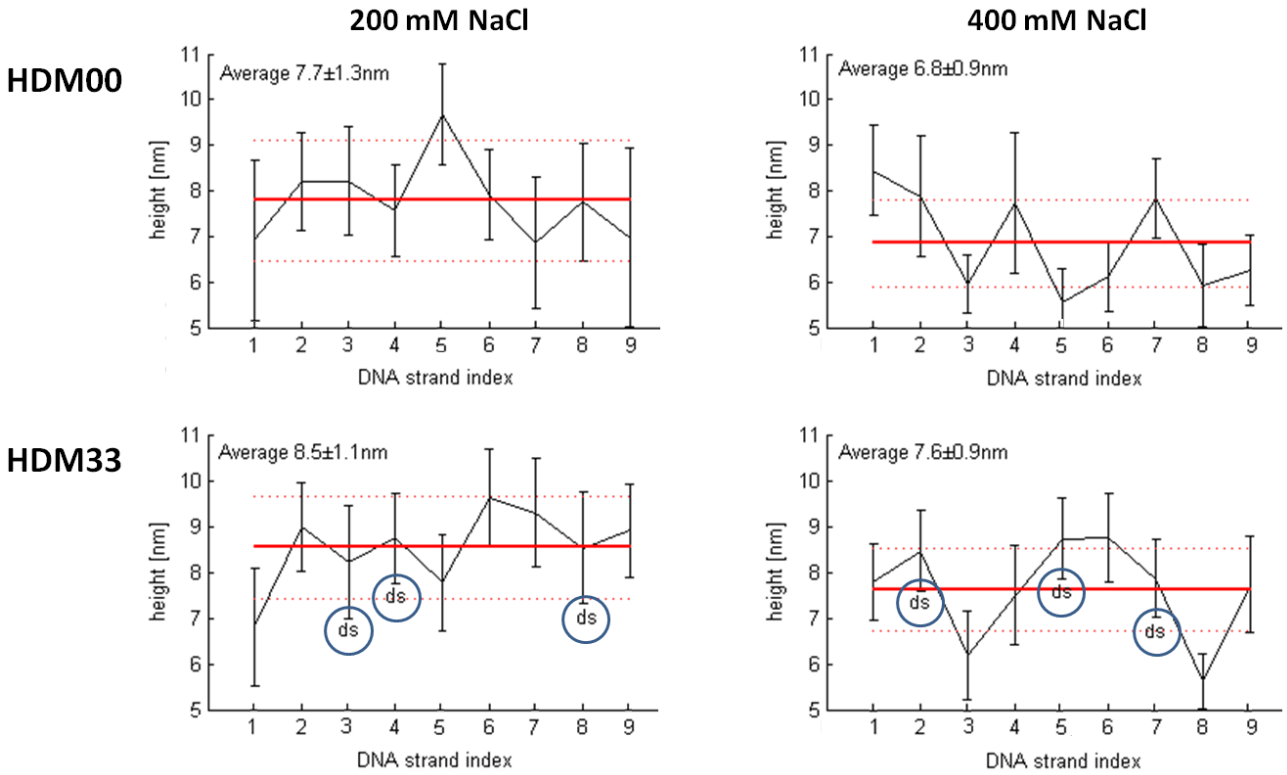

Figure S5. Individual heights of each DNA molecule into the HDM. Heights are measured considering topmost atom of each DNA in the patch respect to the gold surface (z = 0). Average heights are identified by the red lines. dsDNA in the **HDM33** systems are identified by blue circles.

This journal is © The Royal Society of Chemistry 2013

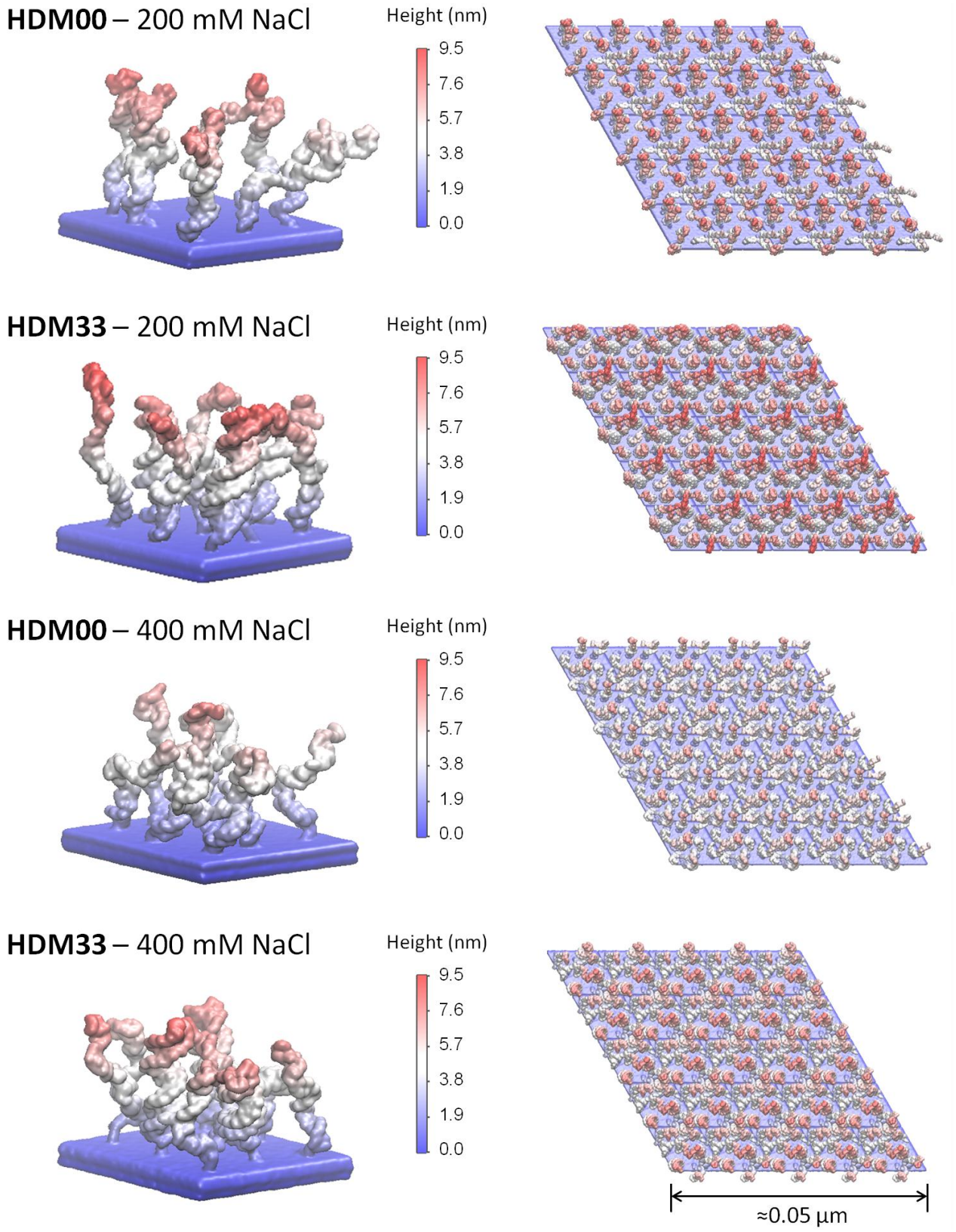

Figure S6. Molecular-level details about the HDMs surface. (a) Snapshots taken from the MD simulation of **HDM00** and **HDM33** at 200 mM and 400 mM NaCl. The patch is represented according to the van der Waals surface and colored according to the height (*z* axis). The simulated

This journal is © The Royal Society of Chemistry 2013

boxes replicated on the *xy* plane – 5 periodic cell copies along *x* and *y* axes for a final system size of ≈0.05 μm. Red zones identify higher regions in the patch.

MD can capture high resolution details about the surface shape and asperities. Figure S3 shows equilibrated MD snapshots of the systems colored according to the patch height (*z*), and the simulated systems replicated on the *xy* plane.

The atomic density expressed as a function of the distance from the surface is known to be fundamental in the characterization of the behavior of DNA monolayers, as well as the absorption of other molecules into the patch.[S19] Indeed, in the framework of polymer brush theory the energy of a DNA monolayer can be expressed as a functional of the this density profile. The array polymer density is calculated self-consistently from the position vectors according to a Hamiltonian that accounts for the conformational energy of polymer chains and the charge interactions. In our approach, DNA molecules are treated explicitly through the positions of all the atoms. In this case the same density can be estimated according to Eq. (S2):

$$\rho(z) = \frac{\Delta t}{S T_{eq}} \sum_{z_i, t_i} \delta\left(z - z_i(t_i)\right) \qquad \text{(S2)}$$

The sum is performed over the whole set of atomic coordinates (only the component orthogonal to the surface is considered here), averaged over the simulation time (T_{eq}) and then divided by the surface area (S). The *z* dimension was discretized with a 1 Å spacing.

Since the densities themselves provide a clear picture of the theoretical distribution of atoms within the patch along the direction orthogonal to the surface, its height can consequently be estimated according to Eq. (S3):[S20]

$$\int_0^{h^*} \rho(z) dz = 0.98 \qquad \text{(S3)}$$

This journal is © The Royal Society of Chemistry 2013

where h* is the distance at which the 98% of all atoms within the DNA array are likely to be found on average. This particular threshold was chosen to reproduce the patch height calculated as the average distance between the Au(111) surface and the topmost base pair of a double strand DNA. This was considered to be a good reference due to dsDNA intrinsic rigidity and reduced bending. The comparison with the experimentally measured height values was found to be satisfactory.

The iso-density surfaces reported in the text (Figure 4) were then obtained by iterating the previously described procedure for each point of the surface – the slab surface was then discretized with a 2 Å spacing along both the principal direction orthogonal to *z* – *i.e.*, the non-orthogonal sides of the simulation box (the angle comprised between the rhombic base sides is 60°). The collection of heights was then interpolated to obtain the density profile shown in figure 4 in the text.

References

S1 Castronovo, M.; Lucesoli, A.; Parisse, P.; Kurnikova, A.; Malhotra, A.; Grassi, M.; Grassi, G.; Scaggiante, B.; Casalis, L.; Scoles, G. *Nat. Commun.* **2011**, *2*, 297.

S2 Bano, F.; Fruk, L.; Sanavio, B.; Glettenberg, M.; Casalis, L.; Niemeyer, C. M.; Scoles, G. *Nano Lett.* **2009**, *9*, 2614-2618.

S3 Mirmomtaz, E.; Castronovo, M.; Grunwald, C.; Bano, F.; Scaini, D.; Ensafi, A. A.; Scoles, G.; Casalis, L. *Nano Lett.* **2008**, *8*, 4134-4139.

S4 Bosco, A.; Bano, F.; Parisse, P.; Casalis, L.; DeSimone, A.; Micheletti, C. *Nanoscale* **2012**, *4*, 1734-1741.

This journal is © The Royal Society of Chemistry 2013

S5 Castronovo, M.; Radovic, S.; Grunwald, C.; Casalis, L.; Moregante, M.; Scoles, G. *Nano Lett.* **2008**, *8*, 4140-4145.

S6 Jorgensen W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.,* **1983**, *79*, 926–35.

S7 Case, D. A.; Darden, T. A.; Cheatham III, T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R. C.; Zhang, W.; Merz, K. M.; Robertson, B..; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Kolossvary, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Liu, J.; Wu, X.; Brozell, S.; Steinbrecher, T.; Gohlke, H.; Cai, Q.; Ye, X.; Wang, J.; Hsieh, M.-J.; Cui, G.; Roe, D.R.; Mathews, D.H.; Seetin, M.G.; Sangui, C.; Babin, V.; Luchko, T.; Gusarov, S.; Kovalenko, A.; Kollman, P. A., AMBER 11. In University of California, San Francisco, **2010**.

S8 (a) Rai, B.; Sathish, P.; Malhotra, C. P.; Pradip; Ayappa, K. G. *Langmuir* **2004**, *20*, 3138-3144; (b) Hautman, J.; Klein, M. L. *J. Chem. Phys.* **1989**, *91*, 4994.

S9 Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.,* **1995**, *117*, 5179-5197.

S10 Perez, A.; Marchan, I.; Svozil, D.; Sponer, J.; Cheatham, T.E.; Laughton, C.A.; Orozco, M. *Biophys. J.*, **2007**, *92*, 3817–3829.

S11 Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.,* **1998**, *98*, 10089-10092.

S12 (a) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys*., **1977**, *23*, 327; (b) Krautler, V.; van Gunsteren, W. F.; Hanenberger, P. H. *J. Comput. Chem*., **2001**, *5*, 501.

S13 Srinivasan, J.; Cheatham, T. E.Cieplak, P.; Kollman, P. A.; Case, D. A. *J. Am. Chem. Soc.* **1998**, *120*, 9401-09.

This journal is © The Royal Society of Chemistry 2013

S14 Jayaram, B.; Sprous, D.; Beveridge, D. L.; *J. Phys. Chem.* **1998**, *102*, 9571-9576.

S15 Sitkoff, D.; Sharp, K. A.; Honig, B. *J. Phys. Chem*. **1994**, *98*, 1978-1988.

S16 Luo, R.; David, L.; Gilson. M. K. *J. Comput. Chem*. **2002**, *23*, 1244-1253.

S17 Sanner, M. F.; Olson, A. J.; Spehner, J. C. *Biopolymers*. **1996**, *38*, 305-20.

S18 Andricioaei, I.; Karplus, M. *J. Chem. Phys.* **2001**, *115*, 6289-92.

S19 Milchev, A.; Egorov, S. A.; Bider, K. *J. Chem. Phys.* **2010**, *132*, 184905.

S20 Crozier, P. S.; Stevens, M. J. *J Chem Phys.* **2003**, *118*, 3855-3860.