Electronic Supplementary Material (ESI) for Nanoscale.

Site accessibility tailors DNA cleavage by restriction enzymes in DNA confined monolayers

Chiara Rotella, Giovanni Doni, Alessandro Bosco, Matteo Castronovo, Alessandro De Vita, Loredana Casalis, Giovanni M. Pavan and Pietro Parisse

Supporting Info

EXPERIMENTAL DETAILS

Preparation of bioresistant ultraflat alkylthiol SAMs

The first step is the preparation of the gold film using the Ulman procedure (mechanical stripping) to obtain a flat and clean gold surface [S1]. A thin gold layer (100 nm) is deposited on clean silicon surface using electron beam evaporation with a rate of 0.1 $nm \cdot s^{-1}$. The good adhesion between gold and silicon is obtained gluing both the substrates using a negative photoresist. The glue is cured at 130°C for 48 hours. After this, to avoid any thermal stresses and have a good attachment between the gold film and the silicon pieces, the samples are left at room temperature for a slow cooling. The ultraflat gold surface is obtained by the separation of the small silicon pieces from the gold-silicon substrate applying slight lateral pressure to the samples. In this way the silicon pieces detach with a clean gold and thin film on the top of the silicon samples. The fresh flat gold surface is incubated all night in a solution of 300 µM TOEG6 HS-(CH_2)11-(O-CH_2-CH_2)6-OH (top-oligo-ethylene-glycol-6) to create a self-assembled monolayer used to nanograft the DNA patches [S2].
All the experiments are performed in liquid cell, to recreate the physiological environment with a XE-100 PARK/PSIA AFM system.

Grafting of DNA sequences

This phase is characterized by the formation of DNA patches into TOEG6 SAM. This was done using a silicon cantilever with a spring constant of 0.5-1.5 N/m. With the nanografting technique in liquid environment is possible to remove the SAM molecules and promote the substitution with the DNA molecules in the solution. In particular it was used a buffer solution containing thiolate ssDNA at 5 and 10µM and 10mM tris, 1mM EDTA, 100mM MgCl2, 100mM NaCl, ph 7 diluted in MilliQ water. The use of Magnesium helps the elongation of the molecules and increase the density of the patches [S3]. After the nanografting process and before the imaging of the surface the sample was washed with a solution of TE 1M NaCl to remove the DNA molecules that were not attached on the gold substrate.

Imaging before and after the incubation with the enzyme

The imaging process was performed in contact mode with a soft silicon cantilever with a spring constant of 0.03 N/m. During the measurements low force is applied to calculate the profile of the nanostructured DNA so as not to compress the real height. All measurements are performed in a TE 1M NaCl solution. This buffer is

used because the TE decreases the degradation process of DNA molecules while the ions in solution, 1M NaCl, allow the height measure without compress or elongate the DNA chains.
After the imaging of the ssDNA it was proceeded with the hybridization of the DNA molecules. The sample was incubated in 300µl solution containing the complementary DNA at 1µM concentration in TE 1M NaCl for 2 hours at room temperature. After the washing and the imaging of the sample it was started the enzymatic reaction. The sample was incubated with a 300 µl solution containing 0.2 units of DpnII/µl in the NEB* buffer (50 mM Bis-Tris-HCl, 100 mM NaCl, 10 mM MgCl2, 1 mM DTT, pH 6) for 3 hours at 37°C. After the reaction the sample was washed with three different solutions: NEB* buffer, a solution contains a detergent able to remove the enzyme on the surface (10mM tris, 1mM EDTA, tween 0.05%, 1M NaCl, ph7) and finally with TE 1M NaCl. The washing process was followed by imaging of the sample in TE 1M NaCl buffer to calculate the height before and after the enzymatic reaction.

ssDNA sequences

In table S1 we report the list of the thiolated ssDNA molecules containing a recognition sequence of 4 bp in red specific for the restriction enzymes used and a fork at the end and the complementary sequence.

Seq1: CONTROL	HS-((CH)2)6-5'-CAAAACAGCAGCAATCCAAACTAGAGACACCCGATTACAAATGC-3'; 5'-GCATTTGTAATCGGGTGTCTCTAGTTTGGATTGCTGCTGTTTTG-3'
Seq2: DPNII Blunt Ended	HS-$(CH_2)_6$-5'-CAAAACAGCAGCAATCCAAAGATCAGACACCCGATTACAAATGC-3' 5'-GCATTTGTAATCGGGTGTCTGATCTTTGGATTGCTGCTGTTTTG-3'
Seq3: HELICASE Forked	5'TTTTTTTTTTTTACGAAAACATCGGGTACGAGGACGAAGACT 3'-$(CH_2)_6$-SH 5'AGTCTTCGTCCTCGTACCCGATGTTTTCGCTTTTTTTTTTTT 3'
Seq4: DPNII Forked	5'-TTTTTTTTTTTTGAGCAAAGATCGGGTAGCAGGACGAAGACT-3'$(CH_2)_6$ -SH 5'-AGTCTTCGTCCTGCTACCCGATCTTTGCTCTTTTTTTTTTTT-3'
Seq5:DPNII Blunt Ended UP	HS-$(CH_2)_6$ 5'-CAAAACAGCAGCAATCCAAACTAGAGACAGATCATTACAAATGC-3' 5'-GCATTTGTAATGATCTGTCTCTAGTTTGGATTGCTGCTGTTTTG-3'

Table S1. List of sequences used in the experiments.

Computational details

The DNA monolayers described in this work (see main paper) were modeled computationally by means of molecular dynamics (MD) simulations. While all-atom molecular models have been used in the past to study a set of analogous systems with higher DNA surface grafting density [S4], in this case the sensibly lower DNA density in the studied systems required the creation of models with increased size, exceeding the capabilities of fully atomistic MD simulations. Coarse grained (CG) potentials for oligomers offer a suitable option for the purpose. In this study, , we chose to employ the 3SPN.2 coarse grained model recently developed by the group of De Pablo [S5]. Herein, each nucleotide is represented by three distinct beads whose interaction are expressed by a potential modelling the intra-molecular interactions with harmonic and periodic angle-dependent terms, commonly employed in most molecular mechanics potential; specific terms are added to account for the aromatic interactions between stacked base pairs and H-bonds between paired complementary bases. The electrostatic charge is located on the bead corresponding to the phosphate group; the value of the charge assigned to each base depends on the ionic concentration in the solution; the screening effect of ions is treated by considering that a fraction of counterions condense around the oligos, thus

decreasing the charge to an effective value calculate through the Manning's Counterion Condensation [S6]. The solvation contribution is described with an implicit solvation term.

The parametrisation of this CG model allowed to reproduce closely a set of DNA experimental measurement such as the melting curves and the hybridisation free energies, while being able to describe the stability of the molecular structures of both B-DNA over the correct range of temperatures and was tested upto microsecond long simulations [S5]. While presenting significant similarity with the earlier Ouldridge's coarse grained potential [S7], this model was preferred since it offers analogous accuracy and a simpler implementation. This implied that the model was fitted using a smaller set of parameters, thus enabling a higher control over the simulations. Moreover, the 3SPN.2 potential is available for in LAMMPS simulation packages [S8], which facilitated the efficient implementation of our CG models.

The monolayer CG models were built as follows. First, the CG model for each distinct oligos was constructed; each oligo was then aligned in a hexagonal arrangement (though as being the most general placement of molecules onto a surface) at a first neighbor distance corresponding to the experimental densities (equivalent to 8 nm of inter-DNA spacing). The 8 aliphatic long linker used to tether the oligos to the Au(111) surface, and the Au(111) surface itself were included implicitly in the CG models since the force field parameters for the latters were missing. For this reason, we fitted a tethering potential applied to the last bases adjacent to the surface using the data from the full atomistic simulations conducted for a previous work [S4] to determine the average distribution of the position and the orientation of the center of mass of the ribose group to which the linker is attached, with respect to the tethering point on the surface. In this way, this additional potential allowed to implicitly accounting for the presence of the aliphatic groups between the first ribose beads and fictitious particles placed on the surface (these particles are effectively not interacting and are just meant to serve as anchoring points kept still by positional restraints). The tethering potential constitutes of three terms: (i) an harmonic potential to maintain the bases at the correct distance from the surface (i.e. the fictitious particles), matching the analogous case for full atomistic simulations; (ii) a three body term to keep the distribution of the angle formed between the perpendicular and distance vector joining the fictitious atom and the ribose bead; (iii) a second three body term to avoid unphysical orientation of oligos with respect to the surface. Since the distributions obtained from our preliminary atomistic simulations showed that the oligos stays at a sufficiently large distance from the Au surface (Figure S1), the presence of the latter has been modeled by a LJ potential, using an attractive and distance parameters as in [S4].

We also built analogous CG models for forked and blunt-ended low density DNA patches, having inter-DNA grafting distance in the hexagonal lattice of 15 nm. These additional systems were used to test more extensively the consistency of our CG models with the experiments. The heights of these additional fork and blunt-ended patches as measured from our CG-MD simulations (Figure S2) were found in good aggreement with the experimental ones measured by AFM, which reinforced our confidence in the CG models.

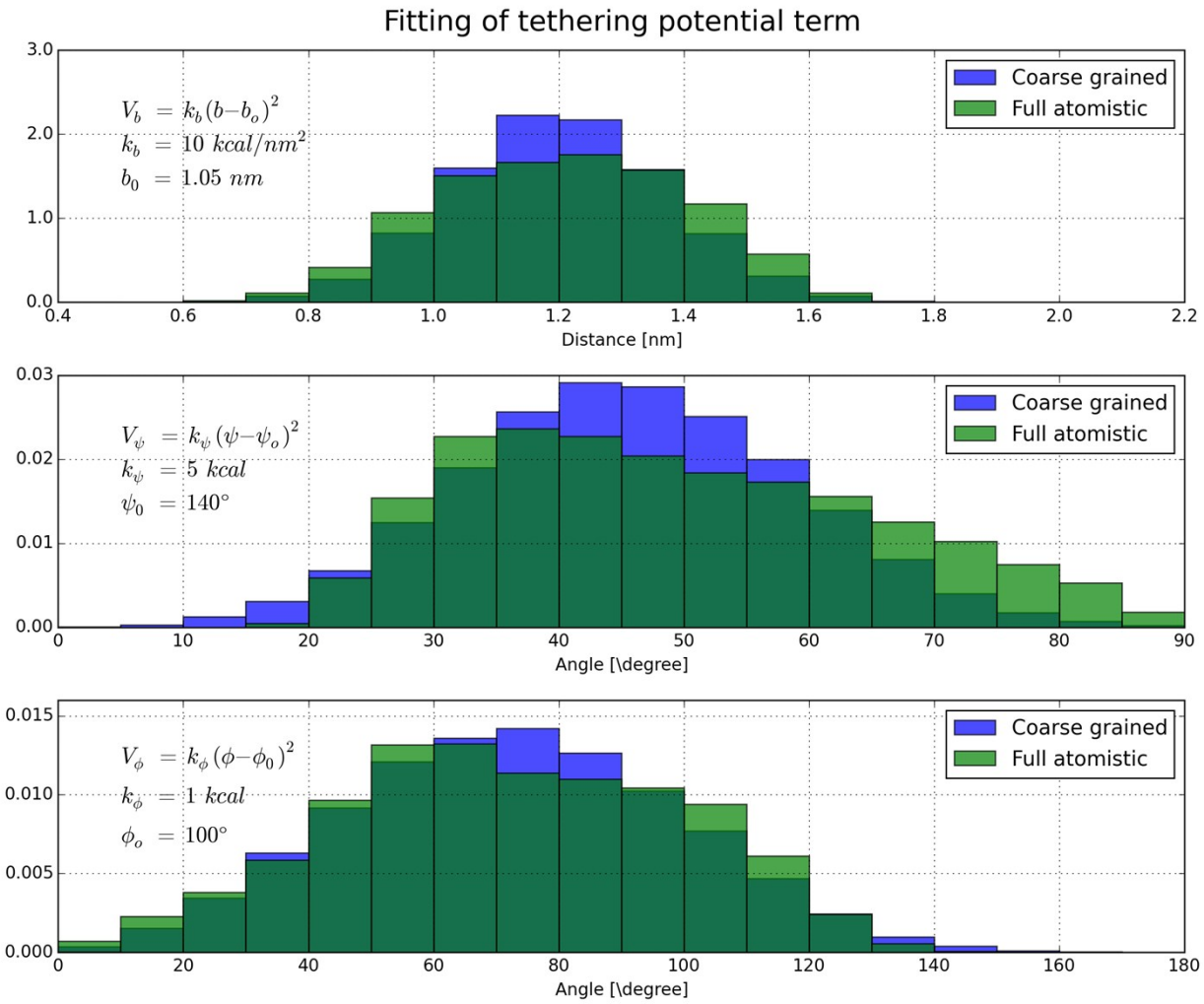

Figure S1: Histograms of the average distributions of geometrical parameters from the full atomistic simulations in [S4] (green) and obtained by fitting the tethering potential in the CG-MD models (blue).

A total of four CG model systems were created for the blunt ended and the forked monolayers, each made of a 10x10 array of oligos arranged according to a hexagonal lattice with inter-DNA grafting distance of 8 and 15 nm respectively. The sequences used are Seq2 and Seq3 reported in table 1.

The counterion condensation effect [S5] is taken into account by the 3SPN.2 force field by rescaling the charge on the oligos to a value of -0.6 *e* corresponding to the effective charge present at 300 K for a linear charge density of about 3.5 A/e. The electrostatic interaction includes solvation contribution implicitly by employing the following potential term:

$$U_{elec} = \sum \frac{q_i q_j e^{-r_{ij}/\lambda_D}}{4\pi\epsilon_0\epsilon(T,C)r_{ij}}$$

were r_{ij} is the separation distance between charges $q_i$and q_j, $\lambda_D$is the Debye screening length of the solution and $\epsilon(T,C)$ is the dielectric permittivity, function of the temperature and the salt concentration. The Debye screening length is defined as follows:

$$\lambda_D = \sqrt{\frac{\epsilon_o \epsilon(T,C)}{2\beta N_A e^2 I}}$$

where β if the inverse of the thermal energy, $N_A$is the Avogadro number and I is the ionic strength of the solution. In the assumption that the contribution to solution of dielectric permittivity of temperature and molarity of NaCl are independent [S6], $\epsilon(T,C)$can be calculated as:

$$\epsilon(T,C) = \alpha(T)\gamma(C)$$

where:

$$\alpha(T) = 249.4 - 0.788T + 7.2 \times 10^{-4}T^2$$

and

$$\alpha(C) = 1.00 - 2.55C + 5.15 \times 10^{-2}C^2 - 0.689 \times 10^{-3}$$

The parameters in the equations reported above and obtained from the fitting of experimental data in [S9, S10] . The values used in the simulation are automatically calculated by the Lammps implementation of 3SPN.2 and assigned to the U_{elec} pair potential.

The system were then simulated by means of MD using LAMMPS [S8], using the Langevin thermostat, with relaxation time of 20 ps; a verlet algorithm was used for integration using a 20 fs timestep. The force cutoff was set to 50 Å, a value that was considered sufficient to avoid long-range corrections. The temperature was progressively increased from 0 to 300 K; the CG systems were then equilibrated for 150 ns of CG-MD simulation, during which the systems reached the convergence of the total energy and of the average height profiles (Figures S2 and S3), while the subsequent 600 ns of the CG-MD runs were used for the analysis.

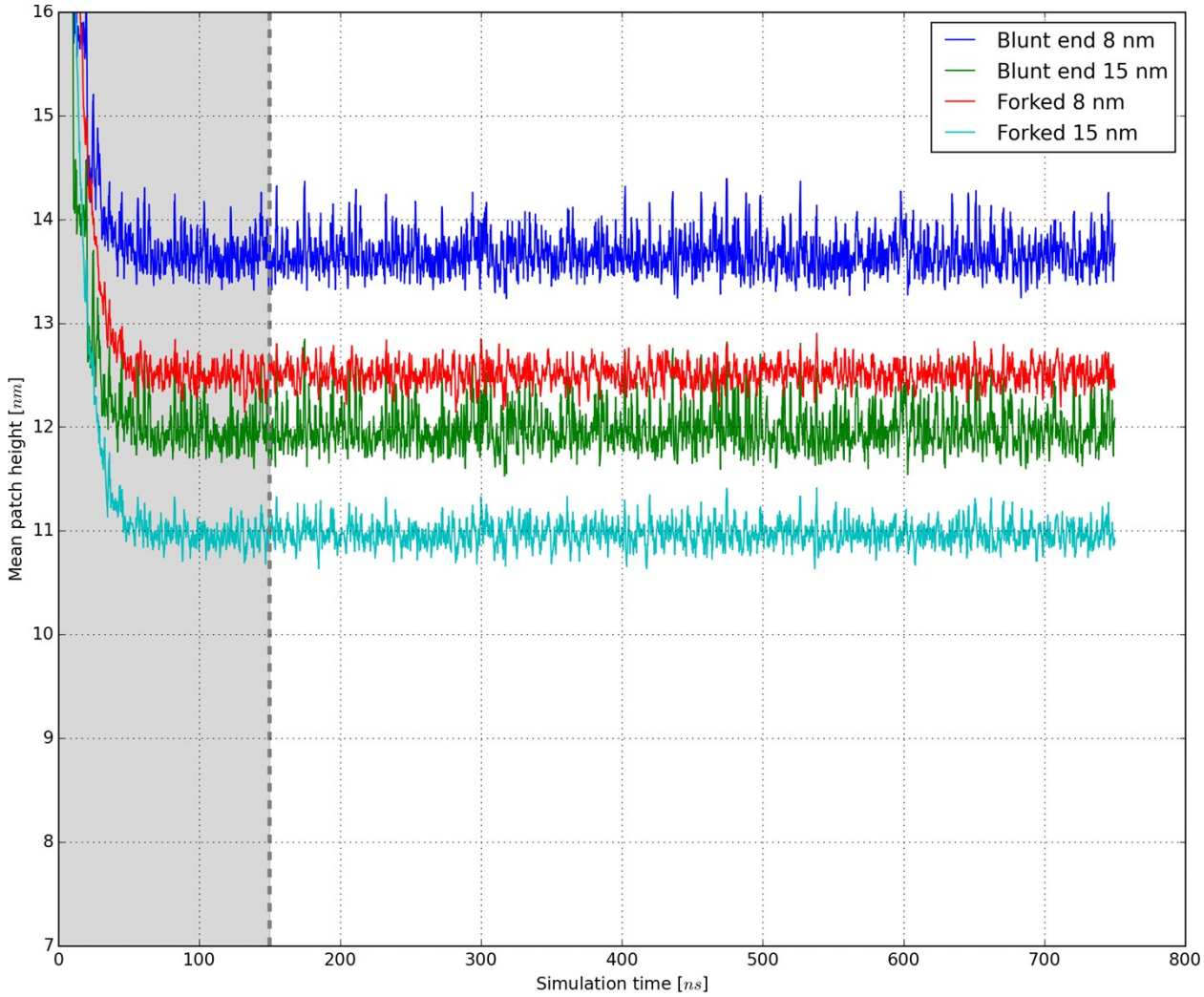

Figure S2: Mean patch heights of the DNA patch for all four CG systems plotted against the CG-MD simulation time; the area shaded in gray refers to the equilibration phase.

The DNA patch densities, and the related enzyme penetration probabilities in the various systems, were calculated as follows. The average DNA planar particle densities were extracted and averaged from the equilibrated phase CG-MD trajectories. In particular, the patch particle density was calculated as a function of the distance from the Au surface every 2 Å of height. The density of the cleavage restriction site as a function of *z* was also calculated according to the same approach.

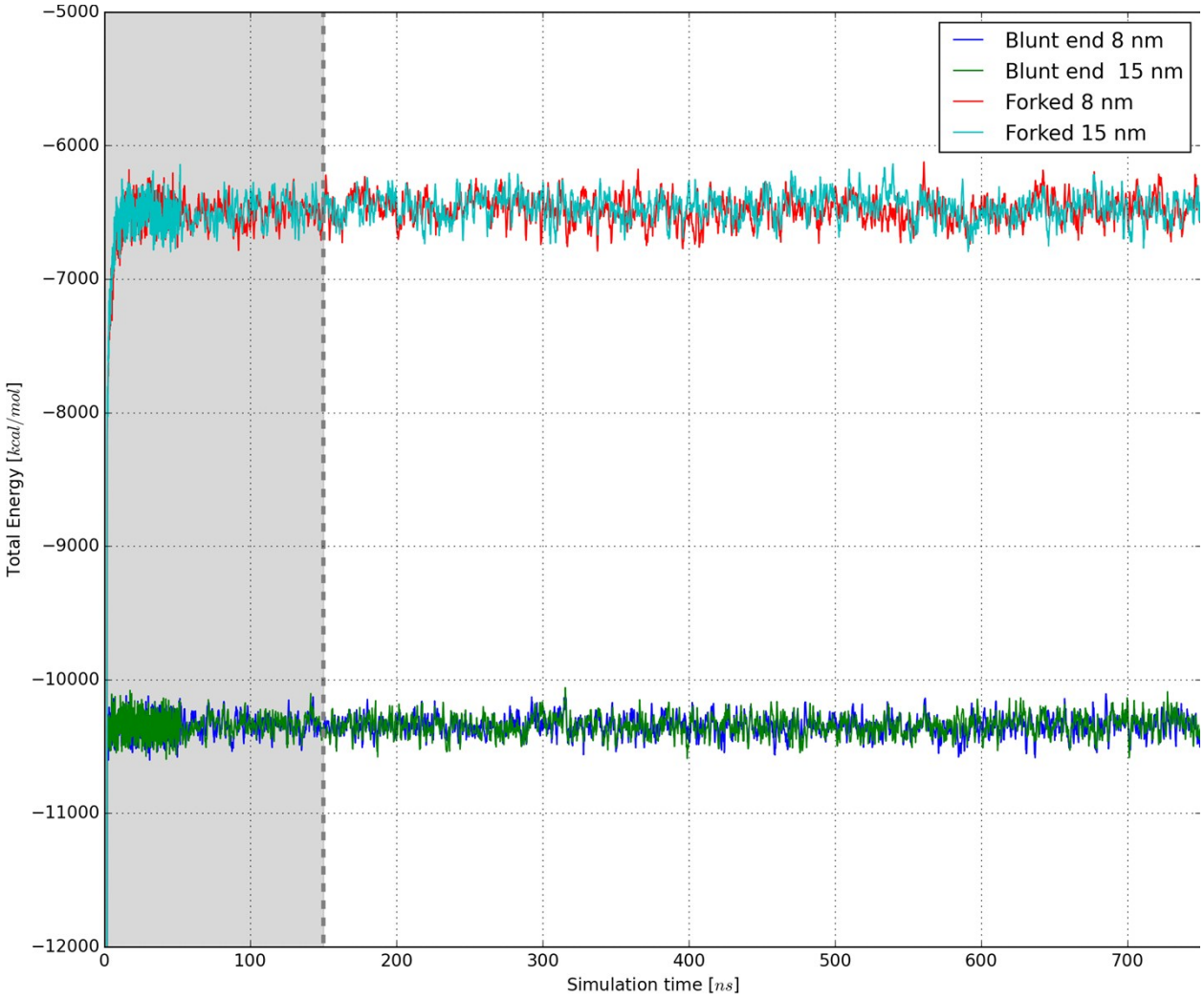

Figure S3: Total energy of the DNA patched plotted for all four CG systems over the CG-MD simulation time; the area shaded in gray refers to the equilibration phase.

The assumption behind the calculation of the penetration probabilities relies on the fact that the penetration of the enzymes is necessarily dependent on the morphology of the patch. In particular, we assumed the latter to be proportional to the void space inside the oligos monolayer effectively accessible by a probe having the same size of the enzyme. The normalised void fraction *ρ(z)* was then calculated by sampling the space let available by oligos in the patch for each each system. This was carried out using a simple Monte Carlo integration algorithm, using a 4.4 nm diameter probe, corresponding in good approximation to the size of the DpnII enzyme (PDB reference: 2DPM).[S11] The *ρ(z)* profiles have been calculated using a 2 Å spacing, which yielded an adequate resolution. Our model represents a bulk 2D monolayer (periodic boundary conditions create an infinite 2D patch) where the enzyme penetrates from the top of the patch toward the Au surface. The densities were then normalised for comparison.

REFERENCES

S1. P. Gupta et al., «Facile Route to Ultraflat SAM-Protected Gold Surfaces by "Amphiphile Splitting",» *Angewandte Chemie International Edition,* vol. 43, n. 4, pp. 520--523, 2004

S.2 R. Levicky et al., «Using self-assembly to control the structure of DNA monolayers on gold: a neutron reflectivity study,» Journal of the American Chemical Society, vol. 120, n. 38, pp. 9787--9792, 1998.

S.3 M. Vidonis, Formation and AFM Characterization of a Magnesium-Based Nucleic Acids Self Assembled Monolayer for Biotechnological Applications, Master Thesis University of Trieste 2012.

S.4 Doni, G. *et al.* Structural and energetic basis for hybridization limits in high-density DNA monolayers. *Nanoscale* **5,** 9988–93 (2013).

S5. Hinckley, D. M., Freeman, G. S., Whitmer, J. K. & De Pablo, J. J. An experimentally-informed coarse-grained 3-site-per-nucleotide model of DNA: Structure, thermodynamics, and dynamics of hybridization. *J. Chem. Phys.* **139,** (2013).

S6. Manning, G. S. Limiting Laws and Counterion Condensation in Polyelectrolyte Solutions. III. An Analysis Based on the Mayer Ionic Solution Theory. *J. Chem. Phys.* **51,** 3249 (1969).

S7. Ouldridge, T. E., Louis, A. A. & Doye, J. P. K. Structural, mechanical, and thermodynamic properties of a coarse-grained DNA model. *J. Chem. Phys.* **134,** 085101 (2011).

S8. Plimpton, S. Fast Parallel Algorithms for Short-Range Molecular Dynamics. *J. Comput. Phys.* **117,** 1–19 (1995).

S9. Stogryn, A. Equations for Calculating the Dielectric Constant of Saline Water. *IEEE Trans. Microw. Theory Tech.* **19,** 733–736 (1971).

S10. Catenaccio, A., Daruich, Y. and Magallanes, C. Temperature dependence of the permittivity of water. *Chem. Phys. Lett.* **367,** 669–671 (2003).

S11. Tran, P.H., Korszun, Z.R., Cerritelli, S., Springhorn, S.S. and Lacks, S.A. Crystal structure of the DpnM DNA adenine methyltransferase from the DpnII restriction system of streptococcus pneumoniae bound to S-adenosylmethionine. *Structure* **6**, 1563-1575 (1998).