This is a repository copy of Site accessibility tailors DNA cleavage by restriction enzymes in DNA confined monolayers.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/116599/

Version: Accepted Version

Article:

https://doi.org/10.1039/C7NR00966F

(c) 2017, The Royal Society of Chemistry. This is an author produced version of a paper published in Nanoscale. Uploaded in accordance with the publisher's self-archiving policy.

Reuse
Unless indicated otherwise, fulltext items are protected by copyright with all rights reserved. The copyright exception in section 29 of the Copyright, Designs and Patents Act 1988 allows the making of a single copy solely for the purpose of non-commercial research or private study within the limits of fair dealing. The publisher or other rights-holder may allow further reproduction and re-use of this version - refer to the White Rose Research Online record for this item. Where records identify the publisher as the copyright holder, users can verify any specific terms of use on the publisher's website.

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.
Site accessibility tailors DNA cleavage by restriction enzymes in DNA confined monolayers

Chiara Rotella a,b,†, Giovanni Doni c, Alessandro Bosco a,†, Matteo Castronovo d, Alessandro De Vita b,c, Loredana Casalis a,e, Giovanni M. Pavan f,* and Pietro Parisse a,e,*

Abstract
Density-tunable nanografted monolayers (NAMs) of short oligonucleotide sequences on gold surfaces show novel properties that make them suitable for advanced biosensing applications, and in particular to study the effects of crowding and confinement on biomolecular interactions. Here, combining Atomic Force Microscopy nanolithography, topography measurements and Coarse Grained Molecular Dynamics simulations, we investigated restriction enzyme reaction mechanisms within confined DNA brushes highlighting the role played by the DNA sequence conformation and restriction site position along the chain, respectively, in determining the accessibility of the enzyme and its consequent cleavage efficiency.

Introduction
The study of enzymatic reactions on nucleic acids in crowded environments has a twofold purpose: helping to understand the processivity of enzymes in synthetic systems, mimicking cell compartments conditions (1-6); and opening up prospects for the development of novel biosensing strategies (7-11). In this respect, DNA processing enzymes reactions within surface-bound nucleic acid brushes has been thoroughly investigated from the early 2000s due to the rapid diffusion of DNA microarray and next generation sequencing technologies, to enable an increasingly effective simultaneous reading of multiple DNA sequences (12-15). However, when entering into contact with surface-bound oligonucleotides, enzymes need overcome side reactions, and steric hindrance due to the crowding of the DNA brushes. Understanding these mechanisms is therefore an essential prerequisite for the optimization of efficient biosensors, especially when miniaturization comes into play.

Recently several reports on restriction enzymes reactions on surface-bound DNA have been published. Different groups have independently proven that restriction enzymes activity can be inhibited by increasing DNA brush density, highlighting the relevance of steric hindrance for the recognition and cleavage of the specific sequences (16-20).

In particular, based on fluorescence and atomic force microscopy (AFM), Castronovo et al. (18, 19) demonstrated that the action of the DPNII enzyme on DNA confined monolayers can be fully inhibited if the DNA packing density exceeds a threshold value that is strictly related to the size of the dimeric form of the DPNII enzyme and that other restriction enzymes (BAMHI and BFAST) also exhibit a similar behaviour. Conversely, Parisse et al. observed no effect of DNA packing density on helicases reactions within similar DNA confined monolayers (21). Such difference has been ascribed to the different mechanisms of action of the two families of enzymes: helicases are molecular ATP-driven motors that can bind and translocate along the DNA sequence, overcoming the steric hindrance of dense DNA brushes, which instead constitutes the major limit to restriction enzymes lateral diffusion. However, the two classes of experiments cited above differed also for what pertains to the choice of the DNA: the forked DNA used in case of helicase reactions could play a major role in facilitating enzyme access from the top-layer side with respect to the blunt-ended DNA used for restriction
enzymes studies. The proper choice of the DNA sequence is then instrumental for the optimization of multiple-reading, multiple-operations miniaturized devices.

In this work, we systematically investigated the effect of DNA conformation on the restriction reaction in different crowding conditions. To this end, by exploiting AFM nanografting, we realized nanoarrays of vertically oriented, thiolated dsDNA molecules of variable density on an ultraflat gold film surface. We chose two different DNA sequences (44 and 42 nucleotides (nt) long, respectively) presenting a 12 bases fork, a variably located restriction site or a combination of the two.

We checked the efficacy of the reaction by monitoring via AFM topography the height variation of the DNA nanostructures, created at variable grafting densities, with respect to an embedding polyethyleneglycol-terminated alkanethiols self-assembled monolayer (SAM), which acts as a reference layer (22). We then rationalized our results by modelling the systems via coarse-grained molecular dynamics simulations (CG-MD). We found that the relative penetration probability of the enzyme into the patch is impacted by the different conformation/functionality of the DNA sequences (blunt ended VS forked) into the patches, providing a molecular rationale for the differences in enzymatic activity seen in the experiments.

Figure 1. (a) Schematic representation of the dsDNA sequences used in the experiments (full sequences are reported in Electronic Supporting Information (ESI). The consensus sequence GATC for the DpnII restriction enzyme is highlighted in red. (b-e) Atomic Force Microscopy images and corresponding line profiles of two nanografted dsDNA patches of DpnII forked sequence of in a T-OEG6 SAMs before (b-d) and after (c-e) the enzymatic reaction.

Results and discussion

Experimental results

As previously reported in other works of from our group (18,19,21), we created via AFM nanografting surface-tethered nanobrushes of dsDNA of different conformation/functionality and different density, on a Au film surface, to assess their impact on surface-confined enzymatic activity. After DNA immobilization, we used AFM topographic measurements and checked the height variation over the patches before and after interaction with the enzyme (Figure 1a, 1b respectively). A significant variation of the height is a sign of an effective action of the enzyme.

Since the density of the monolayer cannot be directly measured, we used the height of the DNA nanostructure before the enzymatic reaction as a reference parameter, which is directly related to the latter: the higher is the packing density of the DNA molecules on the surface, the larger is the resulting nanostructure’s height. This is in fact due to the strong electrostatic repulsion between the negatively charged strands backbones, which causes the DNA molecules to overstretch in the
longitudinal direction. Patch height is precisely monitored through AFM measurements (23,24). To quantify the efficacy of the reaction we then plot the measured ratio between the height of the nanostructure after and before enzyme loading versus the initial height of the patch.

In Figure 2, we resume our previous findings: for blunt ended DNA with a restriction site in the middle of the sequence (Seq2 in Figure1) we observe a threshold density after which the restriction enzyme is not able to cut (Figure 2: blue markers), whereas for the forked DNA (Seq3 in Figure1) the helicases can ubiquitously act for all the densities (Figure 2: green markers).

![Figure 2](image)

**Figure 2.** Final height after enzymatic reaction/initial height ratio versus initial patch height relative to the DpnII reaction, for CONTROL (solid red circles) dsDNA matrices, DPNII (blunt Ended DNA) (solid blue circles) and for Helicase reaction on forked sequences (solid green triangles) (re-adapted from refs 18,21). Data are means +/- s.d.

Here we want to rationalize this effect and to understand if the forked sequence, or more in general, if ssDNA stretches increase the accessibility of protein to DNA specific sequence and favours enzymatic reactions in crowded environment using the DNA brushes as model system. To prove this hypothesis we chose to run DPNII reactions on two different sequences: a forked DNA with the consensus sequence in the middle of the strand (sequence 4 in Figure 1); a blunt-ended DNA with the consensus sequence closer to the liquid-side of the monolayer (sequence 5 in Figure 1), and therefore more accessible for the reaction. Results obtained for these two sequences are summarized in Figure 3.

The normalized ratio between initial and final height after DPNII enzyme exposure on DPNII forked DNA (black full squares in Figure 3 a) is always around 0.67 irrespectively of the initial patch density. The independence of the final patch height on initial dsDNA density suggests that the presence of the fork might favour enzyme access to the enzymatic-cleavable site from the topmost interface of the nanostructures. The softer DNA carpet makes the cleavage site more accessible, *i.e.* more exposed to the enzyme molecules in solution.
To corroborate this result, we performed experiments on the blunt ended duplex DNA molecule with the cleavage site closer to the topmost interface (purple triangles in Figure 3b). Being the site 11 base pairs below the topmost terminus, the expected height reduction ratio of 0.7. As shown in Figure 3,b this is exactly the measured value up to a critical threshold of initial patch height of about 14 nm. This value is 2 nm higher than the threshold evidenced in Figure 2 for a blunt-ended sequence with the cleavage site more buried inside the patch. Considered that these DNA nanostructures possess a certain degree of flexibility in solution (26), this effect could be also rationalized, more in general, with a higher exposure of the consensus site to the solution.

![Figure 3](image)

**Figure 3.** (a) the final dsDNA height after enzymatic reaction/ initial height ratio versus initial height relative to the DpnII reaction, for a forked sequence with a restriction site in the middle (solid black squares) and control (solid red circles) dsDNA matrices. (b) the final dsDNA height after enzymatic reaction/ initial height ratio versus initial height relative to the DpnII reaction, for a blunt ended sequence with the restriction site positioned close to the top of the DNA (open purple triangles) and control (open red circles) dsDNA matrices, and for Helicase reaction on forked sequences (green triangles) . Data are means +/- s.d.

In summary, at experimental level we observed that the DNA brush density is not a crucial parameter for the DPNII reaction in the case of forked sequence, whereas for the blunt-ended sequence the top-access to the nanostructure is critically dependent on the vicinity of the consensus sequence to the solution side, and it is generally inhibited at higher patch densities. These new results broaden the recently reported scheme of two-dimensional lateral diffusion of the enzymes inside the DNA carpet from the side (18). The DNA monolayer should preferably be seen as an ensemble of DNA molecules in continuous interaction with each other, entangling and disentangling during the proceeding of the reaction. Therefore, it is not possible to strictly classify the reaction pathway as penetration from the top or from the side of the patch. It is probably more realistic to imagine that the enzyme will penetrate into the DNA carpet starting cleavage wherever the restriction site is more accessible from the solution.
**Computational analysis**

In order to quantitatively rationalize our results, we performed coarse-grained molecular dynamics (CG-MD) simulations of the DNA patches in the case of blunt ended and forked sequences by using the 3 Sites per Nucleotide (3 Spn2) DNA CG force field developed by De Pablo (25). We built CG models of DNA brushes composed of blunt-ended or fork-ended dsDNAs arranged into a hexagonal planar packing onto a surface and having grafting density higher than the threshold density observed in our experiments. The higher surface density in the CG models corresponds to an inter-strand separation of 8 nm. All DNAs have a restriction site in the middle of the strands (Figure 4: GATC, in red). Figure 4 shows the two simulation boxes containing an array of (10x10) blunt-ended (blue) or forked (grey) DNA CG models initially grafted according to an exagonal lattice with an inter-DNA spacing of 8 nm. The grafted DNA strands replicate in the xy plane through periodic boundary conditions, effectively modelling the bulk of DNA blunt-ended or forked patches (26). Details are available in the Methods section and in the ESI. In the same figure we also display a DpnII enzyme (green), to compare the size of the different molecules in play.

![CG-MD simulation of DNA brushes for blunt ended (a) and forked (b) sequences. In red we evidenced the GATC sequence. In green we show the X-ray structure of the DPNII enzyme for comparison.](image)

At a first-glance analysis of the two screenshots of the boxes, no relevant differences can be observed. The DNA brushes appear homogeneously distributed on the surface, instantaneously exposing “free” surfaces of the dimension of the enzyme in both cases, while in Figure 4 it is evident that the restriction sites could be accessible, to some extent, to an object of the size of the enzyme from the top of the patches at this density. In analogy to our previous paper (26), here we were able to measure the average height of the DNA brushes in the box, thus obtaining 13.6 +/- 0.5 nm for the blunt ended DNA, and 12.5 +/- 0.5 nm for the forked DNA. The difference in height accounts for the ssDNA portions of the forked sequence and is in line with the experimental observations of a maximum height of the nanostructures for the forked-DNA nanostructures systematically lower than that of the blunt-ended ones.
The MD-CG simulations allow us to analyse the molecular crowding inside the brush through the analysis of vertical (z axis) atomic density profiles of the simulated systems, which can be calculated from the equilibrated CG-MD trajectories. In Figure 5 (top), we show the oligonucleotide density for the simulated systems. As expected, a small difference can be observed mostly in the upper part of the patches, where the different structure of two sequences results in a ≈20% lower atomic density. In Figure 5 (bottom panel) the normalized particle density along the vertical direction is shown for the GATC consensus sequence only. As we can see, for both sequences, the restriction site is detected starting from ≈9 nm of height from the surface, while the maximum GATC density is found at about ≈7 nm. Therefore, the restriction site is accessible by the enzyme only if the enzyme is able to penetrate the DNA layer from the top until reaching at least the 9 nm quote.

![Normalized CG density](image)

**Figure 5.** Normalized CG density for the whole sequence (solid lines) and for the GATC portion (dotted lines) of the sequence for the blunt ended and forked sequences obtained from the CG-MD simulations.

Taking into account the size of the enzyme in its dimeric, active configuration, the variation of the normalized DNA density, and the presence of empty space into the two DNA brush models, we estimated the relative probability for the enzyme to penetrate into the two types of DNA patches from the solution (i.e., from the top in the case of a bulk patch model) as a function of the distance from the Au surface. In particular, we discretised the DNA brushes along the patch’s height (dimension orthogonal to the Au surface, z) into layers of 0.2 nm of thickness. For each of these layers, we calculated the probability for a probe having the same size of the enzyme to fit within the void spaces present inside the patch. Then we calculated the progressive penetration probability of the enzyme from the top (from the solution, where enzyme penetration is supposed to start in a bulk patch) to the Au surface (see also the ESI). In Figure 6 reports shows the progressive enzyme penetration probability inside the DNA brushes as a function of the patch height for these high-density systems. Starting from the solution, where the probability is 1, the enzyme penetration probability decreases progressively as penetrating deeper into the brush.
We can see that the curves for the Forked sequence (solid black squares) and for the Blunt Ended sequence (solid blue circles) differ for the heights between ≈6 nm and ≈12 nm. To further highlight this, we plotted the difference (Figure 6: yellow markers) between the access probability for the forked sequence (black) and that of the blunt ended sequence (blue). This analysis shows that the probability of access within the region of the restriction site is higher for the forked sequence than for the blunt-ended one. Notably, this region comprises the zone where the GATC is more likely found (see also Figure 5). Thus, the enzyme appears to have increased probability to penetrate inside the brush from the solution and to reach the restriction sites in the case of the forked sequence than in the blunt-ended one. This can be ascribed to the conformations assumed by the forked or blunt-ended patches in solution, providing a molecular-level rationale to explain the absence of a notable density effect (or less pronounced) for the forked sequences. In general, depending on the sequence chosen, the conformation assumed by the DNAs in solution can favour/hinder the action of the enzyme due to a different crowding inside the patch, consistent with an increased/reduced probability for enzyme penetration inside the DNA matrix.

These theoretical results are also found consistent with additional experimental evidences, showing that shifting the GATC restriction site closer to the patch surface (i.e., to the solution) in the blunt-ended DNA strands produces an increase in the enzyme cleavage efficiency (higher probability for the enzyme to access the restriction site), (see Figure 3b) provoking a rigid shift towards higher densities thresholds for the restriction enzyme with respect with the blunt ended sequence with the restriction site in the middle (Figure 2). In this regard, moving the restriction site closer to the surface in the case of a blunt-ended patch is probably not sufficient to completely overcome the steric hindrance effects, as it is more efficiently accomplished in the case of a forked brush.

**Conclusions**

We used Atomic Force Microscopy-based nanolithography to create confined brushy DNA matrices of different surface densities and we studied the action of restriction enzymes as a function of DNA matrix changes. We find that forked sequences are not significantly affected by surface density,
whereas for blunt ended ones the position of the restriction site causes a rigid shift in the surface density threshold below which the enzyme can operate. These results have been rationalized by using CG-MD simulations of the DNA brushes, obtaining information on the relative penetration probability of the enzyme into the DNA patches from the solution, and on the relative accessibility of the restriction site. This appears to be clearly related to the conformation and flexibility of the two types of DNA patches in solution. The results provide useful insights into the action mechanisms of restriction enzymes in a crowded environment, showing how conformational and structural properties of the DNA can influence enzyme processivity. In particular it seems that the accessibility of a DNA specific sequence in the presence of flanking ssDNA stretches in a crowded environment is increased by ssDNA strands flexibility. Significantly, since in a dense environment DNA molecules behave as a highly dynamic system, in which they continuously entangle/disentangle between them, we highlighted that enzymatic activity cannot be simply predicted taking into account static parameters (i.e. DNA density, position of restriction site, etc) while diffusion mechanisms need to be properly considered. This information is expected to be extremely useful for the optimization of biosensing strategies based on surface-bound DNA matrices but could have implication to understand the effects of DNA conformation and crowding to DNA accessibility also for other DNA binding proteins involved in genome regulation.

Materials and methods
Experimental details
We used Atomic Force Microscopy both to to create the nanobrushes of DNA using the AFM-nanografting technique and to perform topographic measurements of the DNA patches height. For AFM nanografting we used an AFM lever and a liquid cell containing thiolated DNA oligonucleotides to create patches of DNA on a ultraflat gold surface covered by a bio-repellent self-assembled monolayer (SAM) of top-oligo-ethylene-glicol-6 (TOEG6) [22]. The lever scans the surface and applying a force of roughly 100 nN it is possible to displace TOEG6 molecules with the DNA ones creating a nano-assembled monolayers (NAM) of DNA. Changing the number of the lines during the scan process is possible to increase or decrease the density of the DNA patches [24]. The DNA nanobrushes' topography can be then precisely measured using AFM imaging in soft contact mode [23], and the variation of the relative height with respect of TOEG6 carpet before and after the enzymatic reaction can be accurately detected. Extended details on nanografting, imaging and enzymatic reactions can be found in ESI.

Computational Methods
We recently used all-atom (AA) molecular models to study similar DNA monolayer systems with high surface grafting density [26]. However, the low DNA density in the systems studied herein required the creation larger size models, exceeding the capabilities of atomistic molecular dynamics (AA-MD) simulations. Thus, we created coarse-grained (CG) models for the DNA monolayers studied in this work using the 3SPN.2 CG model for DNA recently developed by De Pablo [25]. In this CG model, each nucleotide is represented by three CG beads, while the interaction with the solvent and the effect of ions are treated implicitly (see ESI for details).

We first built CG models for the individual oligos (forked or blunt-ended). Then, consistent with our previously reported AA models for similar systems [26], each CG oligo was aligned in a hexagonal arrangement onto the xy plane at inter-DNA spacing of 8 nm, corresponding to high experimental density. In this way, we produced a 10 x 10 hexagonal array of initially parallel CG blunt-ended or forked DNAs. The aliphatic chains linking the oligos to the Au(111) surface, and the Au(111) surface itself, were included implicitly in the CG models using ad hoc potential terms. The CG models were then refined on the AA ones for the best accuracy (see ESI for details).

The CG models were then simulated by means of CG-MD simulations using the LAMMPS software [27]. All CG-MD simulations used the Langevin thermostat with relaxation time of 20 ps. A verlet algorithm was used
for integration with a 20 fs timestep. All CG-MD runs used 50 Å cutoff, a value considered sufficient to avoid long-range corrections. After an initial heating phase to reach the temperature of 300 K, all models were equilibrated by means of CG-MD simulations. 150 ns of CG-MD were necessary for reaching full equilibration in all cases (convergence of the total energies and of average patch heights profiles, see ESI). After this phase, additional 600 ns of CG-MD were used for data production and analysis. We also built analogous CG models for forked and blunt-ended DNA patches with low density (15 nm of inter-DNA spacing), to test more extensively the consistency of our CG models with the experiments. The average patch heights extracted from the equilibrated phase CG-MD simulations for all fork and blunt-ended simulated systems demonstrated good agreement with the experimental ones measured by AFM. The patch particle density was calculated as a function of the distance from the Au surface (in z-direction). The density of the cleavage restriction site as a function of z was also calculated according to the same approach. The penetration probability for the enzyme from the solution into the DNA patches was calculated from the average density of the patches (ρ(z)) and the void space into the latters (from which the enzyme can penetrate to reach the cleavable sites). This was calculated starting from the topmost of the CG patch models and progressively penetrating in their interior (using a 2 Å spacing). Complete details are provided in the ESI.

Electronic Supplementary Information (ESI) available: Extended details for experimental and computational methods.

Acknowledgements
This work was supported by a grant from the Associazione Italiana per la Ricerca sul Cancro (AIRC) to L. C. and A. B. (AIRC 5 per mille 2011, no. 12214) and a FIRB 2011 grant “Nanotechnological approaches toward tumor theragnostic” (P. L. C.). A.D.V. acknowledges funding from the European Union’s Horizon 2020 research and innovation program under grant agreement No 676580, The NOMAD Laboratory, a Euro-pan Center of Excellence.

Notes and references