



**UNIVERSITÀ DEGLI STUDI DI TRIESTE**

**XXIX CICLO DEL DOTTORATO DI RICERCA IN  
NANOTECNOLOGIE**

---

**DNA NANOBUSHES FOR ENZYMATIC  
REACTIONS, PROTEIN RECOGNITION, AND  
BIOMARKER DETECTION VIA ATOMIC FORCE  
MICROSCOPY**

Settore scientifico-disciplinare: FIS/07

**DOTTORANDA  
ABIMBOLA FEYISARA ADEDEJI**

**COORDINATORE  
PROF. LUCIA PASQUATO**

**SUPERVISORE DI TESI  
PROF. GIACINTO SCOLES**

**TUTORE  
DOTT. MATTEO CASTRONOVO**

**ANNO ACCADEMICO 2015/2016**

**© 2017**  
**Abimbola Feyisara Adedeji**  
**All Rights Reserved**

## Abstract

DNA nanotechnology is an emerging field that provides tools to design and create programmable, spatially-resolved, functional nano-scale devices with potential application ranging from synthetic biology to theranostics.

The first part of this thesis addresses the fundamental question of whether it is possible to study the function and the recognition of proteins in an *in-vitro* milieu that is similar to intracellular conditions, in terms of crowding, confinement and compartmentalization, without the use of crowding agents and obtain information that are triggered by localized and nanoscale crowding. To this end, we design and construct surface-bound functional, sequence-dependent DNA nanoreactors with distinct inherent heterogeneity and homogeneity in their molecular structure, using two approaches: an atomic force nanolithography (AFM)-based technique termed nanografting and the spontaneous formation of self-assembled monolayers of pre-hybridized double stranded (ds)DNA molecules. These nanoreactors then serve as test-bed for understanding the functions and interactions of DNA with protein. With such a tool, we studied the recognition and the specificity of a highly stringent, site-specific endonuclease called BamHI, an enzyme with high specificity for its cognate site and well-defined crystallography structure that is routinely used as protein model and for cloning engineering. Our results show an unprecedented digital behaviour of BamHI triggered by nanoscale crowding within highly dense non-cognate spatially resolved DNA nanostructure.

The second part of this thesis aims at developing a computation-based strategy to optimise cyclic peptides with inherent affinity and selectivity that mimic those of a single loop of an antibody ( $\mu\text{M}$  range affinity). Using a combination of molecular dynamics simulations and Monte Carlo algorithm (tagged algorithm 1), our theoretical collaborators have improved the binding affinities of peptides designed to recognize Beta-2-microglobulin ( $\beta\text{2m}$ ), a biomarker over-expressed in ovarian cancer type 1 (OVA 1). The affinity of the designed peptide to the target of interest is further characterized by developing a label-free experimental validation, which utilizes DNA-directed immobilization (DDI) combined with nanografting to form nanopatches of spatially-oriented DNA-cyclic peptide able to recognize the solvent-exposed binding site on  $\beta\text{2m}$ . Using AFM height measurement, we obtained a micro-molar range binding affinity towards  $\beta\text{2m}$ . Subsequently, our theoretical collaborators, Dr Miguel Soler and Dr Sara Fortuna, employed a new computer-

based protocol, which allows the identification of possible binding sites on  $\beta 2m$  (tagged algorithm 2). Following the generation of peptides, they screen the peptides based on their predicted binding energies, distance between  $\beta 2m$  and peptides, and the hydrophobic surface areas. With such computation strategies, pep331 was selected for experimental validation. Afterwards, we selected one peptide (pep381) from the pool of peptides designed by our collaborators using algorithm 1, and the pep331 generated using algorithm 2, and we set up an AFM based experimental validation, to determine the sensitivity of both peptides. Our results show that both peptides bind  $\beta 2m$  with sensitivity of  $\sim 7 \mu M$ . These results give us a benchmark preliminary data to design surface-bound synthetic bidentate peptides with possible application in biomarker detection and discovery.

## Riassunto (Abstract in Italian)

Il settore innovativo delle nanotecnologie basate sul DNA fornisce nuovi strumenti per la progettazione di nano-dispositivi con specifiche caratteristiche strutturali che possano essere applicati in vari campi, dalla biologia di sintesi alla teranostica.

La prima parte di questa tesi mira a comprendere la possibilità di studiare *in-vitro* come la funzione delle proteine e le loro interazioni sono influenzate dalla variazione su nanoscala della densità molecolare locale, ricreando un ambiente simile a quello intracellulare dal punto di vista della densità molecolare e della compartimentalizzazione spaziale, ma anche dal punto di vista del confinamento delle proteine stesse mediante l'impiego di agenti di affollamento. Pertanto abbiamo creato dei nano-reattori composti da molecole di DNA immobilizzate su superfici con caratteristiche strutturali modulabili mediante l'impiego di diverse sequenze di DNA e sfruttando due approcci di immobilizzazione: da una parte, una tecnica litografica chiamata *nanografting* basata sull'utilizzo del microscopio a forza atomica (AFM), e dall'altra, la formazione spontanea di singoli strati auto assemblati di molecole di DNA a doppio filamento. Il nano-reattore così costituito funge da ambiente per testare e comprendere le funzioni delle proteine e la loro interazione con il DNA. In questo modo abbiamo studiato la specificità e il meccanismo di riconoscimento sul DNA di BamHI, una endonucleasi sito-specifica, che taglia cioè il doppio filamento di DNA, di cui è nota la struttura cristallografica e che viene utilizzata frequentemente come strumento per il clonaggio di proteine. I nostri risultati mostrano per la prima volta come molecole di DNA disposte in maniera predefinita all'interno del nano-reattore, e non affini all'enzima, ne modulino la capacità di riconoscere il DNA sulla base della loro densità.

La seconda parte della tesi mira allo sviluppo di una tecnica computazionale per l'ottimizzazione di peptidi ciclici con affinità e selettività simili a quelle di un singolo ciclo di un anticorpo (affinità micromolare). Usando una combinazione di simulazioni basate su dinamica molecolare e algoritmo di Monte Carlo (algoritmo 1), i nostri collaboratori, esperti nello studio teorico di questi sistemi, hanno migliorato l'affinità di legame tra un peptide e la Beta-2 microglobulina ( $\beta 2m$ ), un biomarcatore over-espresso nel cancro ovarico di tipo 1 (OVA 1). I peptidi progettati per il bersaglio di interesse sono stati successivamente caratterizzati tramite una tecnica sperimentale che non richiede il marcamento del bersaglio (*label-free*), che utilizza l'immobilizzazione diretta del DNA (DDI) combinata con il

*nanografting* per formare, su scala nanometrica, aree di peptidi ciclici legati al DNA. In tal modo, vengono ottenuti peptidi spazialmente orientati e in grado di riconoscere un sito di legame esposto al solvente su  $\beta 2m$ . Con l'AFM, usando misure di altezza, abbiamo ottenuto una serie affinità di legame micro-molari verso  $\beta 2m$ . Successivamente, i nostri collaboratori in ambito teorico, il dott. Miguel Soler e la dott.ssa Sara Fortuna, hanno sviluppato un nuovo protocollo che consente l'identificazione di possibili siti di legame su  $\beta 2m$  (algoritmo 2). Dopo la generazione di peptidi, hanno selezionato una serie di peptidi in funzione delle loro energie di legame, la distanza tra  $\beta 2m$  e peptidi, e le superfici idrofobiche. Con tali strategie di calcolo, pep331 è stato selezionato per la validazione sperimentale. In seguito, abbiamo selezionato un peptide (pep381) dall'insieme dei peptidi disegnati dai nostri collaboratori utilizzando l'algoritmo 1, e un peptide (pep331) generato mediante l'utilizzo dell'algoritmo 2, per implementare una validazione sperimentale basata sull'AFM al fine di determinare la sensibilità di entrambi i peptidi. I nostri risultati mostrano che entrambi i peptidi si legano a  $\beta 2m$  con sensibilità di  $\sim 7 \mu M$ . Questi dati forniscono una serie di elementi basilari di riferimento per la progettazione di peptidi sintetici bidentati da applicare per il rilevamento di biomarcatori su superfici.

## Acknowledgements

Firstly, I would like to express my sincere gratitude to my supervisor, Prof. Giacinto Scoles, for his immense support during my PhD program in Italy. Thank you for inviting me to Italy; it has been a productive journey, that invitation has turned me into who I am today. Thanks for encouraging my research and for giving me an enabling environment to grow as a research scientist. To Prof's wife, Giok Lan Scoles, thanks so much for being there and for supporting my work by assisting Prof, whenever he needs to come to Udine and supervise my experimental work.

Besides my supervisor, I would like to thank my tutor and mentor, Dr Matteo Castronovo, words will fail me if I must recount all your immense contribution towards the making of my research career in Italy, I truly appreciate your efforts in terms of time, ideas, motivation, tutelage, and immense knowledge towards my Ph.D. work. I must say that your passion and enthusiasm are contagious, I am glad our path crossed, thanks for sowing a part of your knowledge and experience into my career. I will never forget your attitude towards research and your spirit of not given up until you achieve the best and these have become my qualities as a young researcher.

My sincere thanks go to Dr Sara Fortuna and Dr Miguel Soler for introducing me to research areas of biomimetic computationally-designed peptides. Your contributions toward my experimental work on spatially confined peptide arrays are highly appreciated, thanks so much, your effort is highly commendable.

Likewise, I was opportune to work with an amazing set of scientists, Dr Alessia Amodio and Prof. Francesco Ricci, thanks for involving me on the work that demonstrates the finely-control of supramolecular assembly of DNA tiles via pH, you both taught me simple but effectual science.

To the members of MONOLISA group in Udine, both past and present, especially, the members of DNA nanotechnology and AFM nanolithography group; Alex Stopar, Lucia Coral, Marco Vidonis, Iman Abdollahzadeh, and Luca Crevatin, thanks so much for your love, openness, and collaborative spirit, you all made my stay in the Lab educative and interesting.

Special thanks go to Lisa Barbiani, administrative secretary of the group, thank you so much dear, for all the assistance that you rendered, your act of kindness can never be erased from my heart.

From my journey before Italy, I am grateful to Prof. Wole Soboyejo, Prof. Douglas. J. Buttrey, Dr Omololu Akin-Ojo, Dr Peter Ngene and Dr Kola Lawal—for your guidance, support, and encouragement—you all believe in me and gave me a chance to learn and harness the opportunities in science and research, thanks so much.

Finally, I thank God for providing energy and strength throughout my PhD program, father you ensure that I excel and that I do not quit even when the journey was becoming tougher, rather you rekindled me and energised me to continue till the end--thank you Jesus. Also to my wonderful and precious family; firstly, I would like to thank my husband, Sam, I love you so much, thanks for your encouragement, unlimited support, sacrifice and your unwavering love, your presence in Italy makes my final year unforgettable, you are the best; Abolaji, I am forever grateful for your support and your believe in me, you are always optimistic and you keep telling me “it will all work out for your good”, thanks my brother; Lastly, to my parent, thanks so much for your prayers, support, understanding and encouragement, the Lord will keep you to eat the fruit of your labour in Jesus name (amen).

## Abbreviations list

DNA	Deoxyribonucleic acid
SAM	Self-assembled monolayer
ssDNA	Single stranded deoxyribonucleic acid
dsDNA	Double stranded deoxyribonucleic acid
PEG	Polyethylene glycol
DPN	Dip-pen nanolithography
MPTMS	3-mercaptopropyltrimethoxysilane
BamHI	from <i>Bacillus amyloliquefaciens</i>
AFM	Atomic force microscopy
DFT	Density functional theory
GIXD	Grazing incidence x-ray diffraction
NHS	N-Hydroxysuccinimide
MCH	6-mercapto-1-hexanol
XPS	X-ray photoelectron spectroscopy
NEXAFS	Near-edge x-ray absorption fine structure spectroscopy
nt	Nucleotide
S/A	Surface density parameter
DDI	DNA-directed immobilization
FDN	Functional DNA nanoreactors
bp	Base pair
TOEG	Top oligo-ethylene-glycol
$\beta$ 2m	Beta-2-Microglobulin
MD	Molecular Dynamics
SPR	Surface Plasmon Resonance
NMR	Nuclear Magnetic Resonance

# Contents

<b>Abstract</b> .....	<b>ii</b>
<b>Acknowledgements</b> .....	<b>vi</b>
<b>Abbreviation list</b> .....	<b>viii</b>
<b>Contents</b> .....	<b>ix</b>
<b>1.0 Introduction</b> .....	<b>1</b>
1.1 DNA and its fundamentals .....	1
1.2 Surface-bound DNA self-assembling nanostructures .....	3
1.3 Structure—function relationship in surface-bound DNA nanostructures; Effect of confinement, crowding, and ionic strength on the DNA conformation and DNA hybridization .....	11
1.4 Synthetic DNA compartments for biochemical reactions.....	20
1.5 Translation of DNA arrays into peptide chips for biomarker detection and protein recognition .....	24
1.6 Thesis outline and contributions .....	28
1.7 References.....	32
<b>2.0 Results and discussion on endonuclease recognition within highly dense DNA nanoreactors</b> .....	<b>42</b>
2.1 Endonuclease enzymes.....	42
2.2 Design and construction of surface-bound DNA nanoreactors .....	43
2.3 Specificity of BamHI reaction within sequence-dependent DNA nanoreactors .....	44
2.4 Sequence-density related action of BamHI within DNA nanoreactors .....	49
2.5 Tracking and capturing the endonucleases recognition within DNA nanoreactors .....	54
2.6 Discussion and conclusion .....	60
2.7 Materials and methods .....	61
Materials.....	61
DNA-sequences design .....	62
Preparation of ultra-flat gold substrate.....	62
Preparation of top-oligo-ethylene-glycol SAM on ultra-flat gold substrate .....	63

	Approach 1: Generation of DNA nanoreactors via nanografting .....	63
	Hybridization of thiol-modified ssDNA nanoreactors.....	63
	AFM nanolithography and imaging.....	64
	Approach 2: Self-assembled monolayer formation of double stranded DNA .....	64
	AFM measurements .....	65
	Restriction enzymatic reaction.....	65
2.8	References .....	66
<b>3.0</b>	<b>Results and discussion on AFM-based characterization and determination of binding affinity of <i>in silico</i> optimized peptides towards a biomarker-protein..</b>	<b>68</b>
3.1	Computation-based design, generation, and screening of cyclic peptides..	69
3.2	AFM-based experimental validation of cyclic peptides 461 towards $\beta$ 2m.	72
3.2.1	Generation of peptide arrays via nanografting and DNA-directed immobilization .....	73
3.2.2	Optimization of the surface density of DNA arrays.....	73
3.3	Functionality of DNA-pep461 nanoassemblages and determination of dissociation constant by Atomic Force Microscopy (AFM).....	78
3.4	Conclusion.....	82
3.5	References .....	84
<b>4.0</b>	<b>Results and discussion on surface-bound high affinity binders for protein recognitions .....</b>	<b>86</b>
4.1	Binding site identification, peptide design, and screening.....	86
4.2	Experimental validation of $\beta$ 2m-peptide binding via atomic force microscopy .....	88
4.3	Discussion and conclusion .....	94
4.4	Materials and methods for the experimental validations for pep461, pep331, and pep381 .....	95
	Instrumentation .....	95
	DNA and peptide sequences .....	95
	DNA-peptide conjugates preparation.....	96
	Preparation of ultra-flat gold substrate .....	96

	Preparation of top-oligo-ethylene-glycol SAM (TOEGSAM) on the ultra-flat gold substrate .....	97
	Nanografting of thiol-modified ssDNA in contact mode.....	97
	Hybridization with complementary DNA-peptide conjugates.....	97
	Beta-2-microglobulin binding assay .....	98
	AFM imaging of DNA-peptide assemblages before and after the $\beta$ 2m recognition assay .....	98
4.5	References .....	99
<b>5.0</b>	<b>Conclusions and outlook.....</b>	<b>101</b>
5.1	Future perspectives.....	101
5.1.1	Outlook on enzymatic reactions within DNA nanoreactors .....	102
5.1.2	Future works on DNA-peptide arrays for biomarker detection ....	102
5.2	References .....	103
	<b>Appendix A: pH controlled assembly of DNA tiles.....</b>	<b>104</b>
A.1	Introduction.....	10
A.2	Results and discussion.....	105
A.3	Conclusion.....	110
A.4	Future perspectives.....	111
A.5	Materials and methods .....	112
	Materials.....	112
	Protected tile.....	113
	pH-dependent circuit.....	113
	Buffer conditions.....	113
	Protected tile annealing .....	114
	Substrates preparation .....	114
	Fluorescence measurements.....	114
	pH-dependent self-assembly of DNA tiles .....	114
	Atomic force microscopy.....	114
	Fluorescence microscopy .....	115
	Measurement of tile-assembly yield .....	115
A.6	Supporting figures.....	116
A.7	References .....	126

<b>Appendix B: SPR, NMR, and ITC characterization of peptides for chapter 3.....</b>	<b>128</b>
B.1    SPR results .....	128
B.2    ITC results.....	130
B.3    NMR results .....	131
B.4    Additional results from computation-based screening.....	134
B.5    Methods for computation-based design .....	135
Peptide evolution.....	135
Molecular dynamics simulations.....	136
B.6    Materials and methods for SPR, ITC, and NMR characterization .....	137
SPR .....	137
ITC .....	137
NMR .....	137
B.7    References .....	138
<b>Appendix C: Computational details for the generation and screening of the peptides for chapter 4.....</b>	<b>139</b>
C.1    Binding site identification.....	139
C.2    Peptide screening: simulation details.....	141
C.3    Peptides screening: computational results .....	141
C.4    References .....	146
<b>List of figures.....</b>	<b>147</b>
<b>List of tables.....</b>	<b>154</b>

## 1.0 Introduction

The inherent capacity of nucleic acids to self-assemble with Watson-Crick base pairing into programmable nanostructures allows their application as promising biomaterials in different research areas. DNA nanotechnology for example has spurred the development of a molecular approach to information technology termed DNA computing, which has led to implications for synthetic biology and nanomedicine.

In this chapter, we outline intrinsic physiochemical properties and biological operations associated with DNA molecules, which make them indispensable building blocks for programmable and innovative functional nano-devices. Also, we report a literature survey of existing immobilization mechanism and nanolithography techniques that are currently used to design and create both microscale and confined nanoscale surface bound DNA nanostructures. In addition, we outline structure—function relationship associated with nano-confined surface bound DNA nanostructures with respect to DNA hybridization, conformation, and mechanical properties.

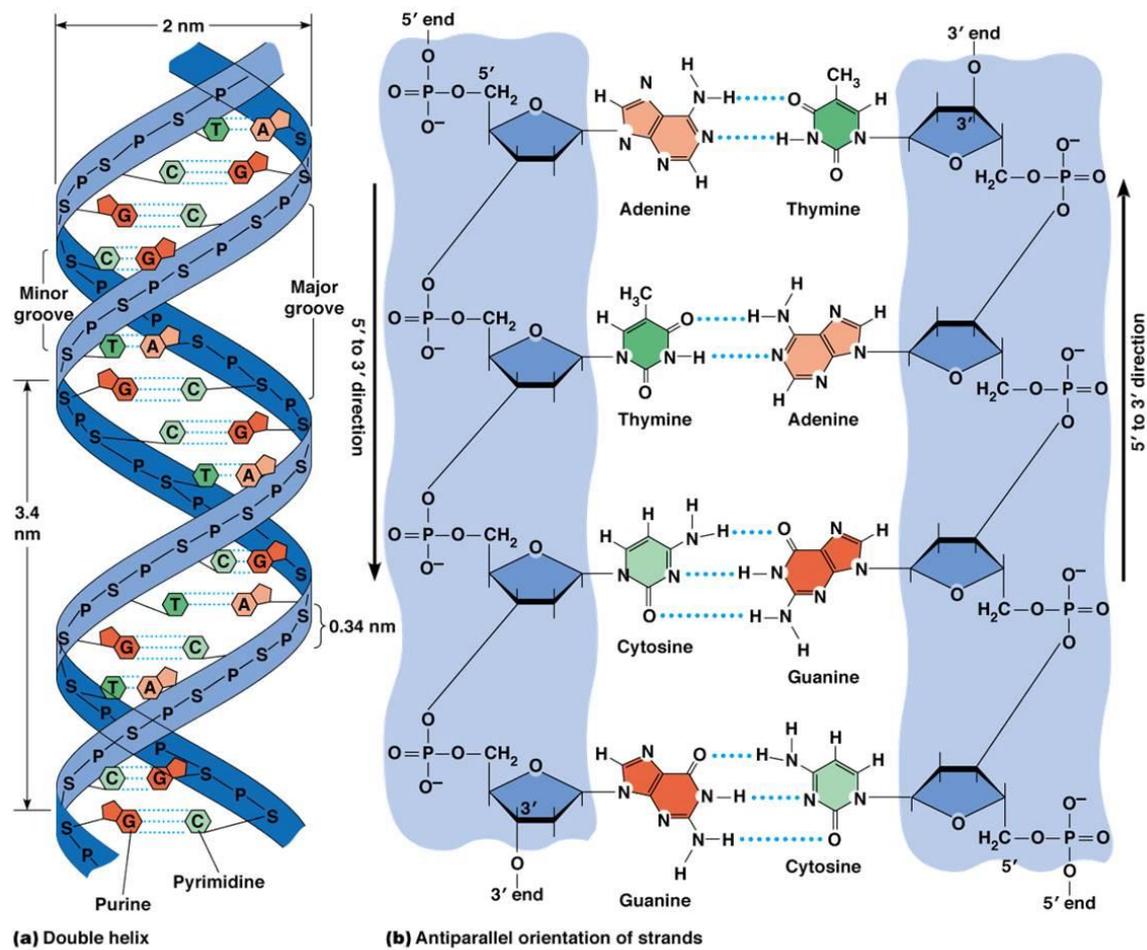
Furthermore, we highlight the emerging application of surface-bound DNA nanostructures in synthetic biology, as it provides the possibility of creating synthetic cell-free, spatially organized, and DNA compartments of different densities, that serve as test-bed to study biochemical reaction in complex constructs akin to the intracellular environment.

Lastly, we outline the contribution of this thesis to the development of functional surface-bound DNA nanostructures with potential application in synthetic biology and protein diagnostics.

### 1.1 DNA and its fundamentals

DNA (2'-deoxyribonucleic acid) is one of the most important biomolecules in living organism because it is capable of forming intrinsic molecular codes that store generic information, which is necessary for growth, development, function and reproduction of nearly all living organisms<sup>1-4</sup>. It is bio-polymeric in nature; consisting of repeats of monomers known as nucleotide group with two polynucleotides chains coiled around the same axis like a double helix<sup>4,5</sup> as shown in Fig. 1.0a-b . Each of the nucleotide is constituted by a nucleoside, pentose sugar and a phosphate group<sup>1,4</sup>.

Additionally, the pentose sugar and the phosphate group are located at the backbone of each strand of the DNA (see Fig. 1.0). The nucleotides can be categorized into two classes: pyrimidines and purines<sup>2,5</sup>. The purines are adenine and guanine while the pyrimidines are the cytosine and the thymine<sup>1</sup>. Also, the nucleotides on the opposite strands interact via the non-covalent hydrogen bonding--adenine (A) forming two hydrogen bonding with thymine (T), while cytosine (C) form three hydrogen bonding with guanine (G)<sup>2,4-5</sup>, as shown in Fig. 1.0b. This interactions is known to be one of the most robust and highly programmable molecular recognition in biology, it is called Watson-Cricks hydrogen bonding base pairing<sup>2-3,5-6</sup>, which is the dominant type of recognition in DNA hybridization.



**Figure 1.0:** Depicts the double structure of B-DNA with vital segments like minor and major groove that occur where the back bones are closer and far apart respectively. The theoretical diameter of DNA double helix (double stranded DNA (dsDNA)) is 2 nm<sup>2,7</sup> and each nucleotide is 0.34 nm long as shown in (a). The chemical and molecular compositions of each nucleotide in (a) are clearly shown in (b) with hydrogen bonding linking each complementary base pair and the antiparallel

direction of each of the complementary helical strand. Adapted from Pearson education incorporation<sup>8</sup>. Copyright © (2012) Pearson education Inc.

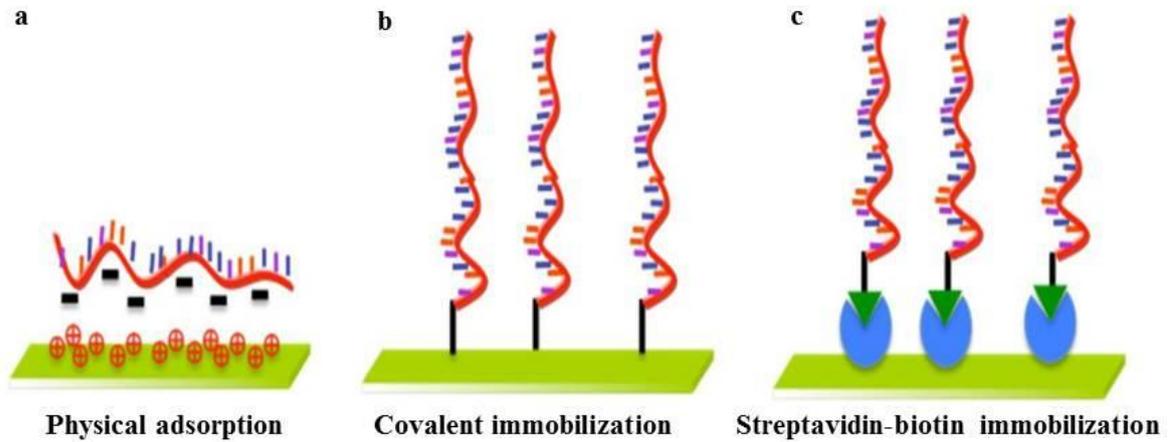
Moreover, the physiochemical properties and mechanical properties of DNA are invaluablely important. For example, the covalent interactions within the atoms and molecules in each nucleotide<sup>1,2,4</sup>, sequence-specific hydrogen bonding within the opposite nucleotides<sup>2,3,5,9</sup>, stacking interactions within the adjacent bases on the same strand<sup>2,4</sup>, high surface charge density across the phosphate backbones<sup>2,3</sup> and the reversible capacity for the double helix to unwind into two complementary strands and vice-versa under the appropriate physiochemical conditions (pH, salts and temperature). The latter processes are termed denaturation and hybridization respectively<sup>2,3</sup>. In addition to the physical properties are the following mechanical properties; mechanical toughness, flexural<sup>2,3</sup> properties with bending and compactness without fracture, all these aid the packaging of DNA in eukaryotic cells<sup>2,10</sup>.

The aforementioned properties made DNA an unprecedented construction material in nanotechnology and nanoscience. DNA nanotechnology is a branch of nanotechnology that is pivoted on the non-covalent interactions within the linear DNA and its robust molecular recognition to construct a programmable functional nano-device of well-defined patterns and shapes<sup>6,11-13</sup>, with potential applications ranging from information technology to synthetic biology<sup>13-16</sup>. Likewise, the significant progress in this field has been ably supported by commercial availability of synthetic DNA with wide range of modification which improves its functionalities<sup>17,18</sup>, existing tool box made up of enzymes and proteins for DNA manipulation<sup>19</sup>, and experimental techniques for characterization and analysis of DNA nanostructures<sup>7</sup>.

## **1.2 Surface-bound DNA self-assembling nanostructures**

One of the leading direction in developing DNA nanotechnology, is by immobilization of DNA molecules on a solid surface, to form locally confined<sup>13,20-22</sup> and non-locally confined DNA nanostructures<sup>23-27</sup> with nano-scale precision and well-defined orientations. This approach is driven by surface chemistry, non-covalent inter and intra molecular forces that aid the self-assembling process, salt concentration in the buffering solution for self-assembling process<sup>24,28</sup>, chemical photolithography<sup>29</sup> and the scanning probe nanolithography process<sup>30,31</sup>.

Surface chemistry plays an important role in immobilizing of DNA molecules on surfaces, and it can be subdivided into two important mechanisms; 1) Non-covalent mechanism i.e. physical adsorption, electrostatic interactions, Vander der Waals interactions and hydrophobic interactions; 2) Covalent and chemisorption immobilization.



**Figure 1.1:** The schematic depicts three different types of immobilization of DNA molecules on solid surfaces that are regularly used in DNA nanotechnology field <sup>32</sup>.

Physical adsorption is a fast and easy approach to immobilize DNA on to a solid surfaces without any modification— it simply employs electrostatic interactions between the cation-coated substrate (mica substrate) and the negative charges located at the phosphate backbone of the DNA molecules <sup>31-33</sup>. This approach has been widely used in development of DNA microarrays and adsorption of DNA origami on cation-mediated surfaces <sup>34,35</sup>, however, it has some limitations ; random adsorption of DNA on surfaces, desorption of the DNA molecules resulting in short life times, and heterogeneous orientation of the DNA on the surfaces<sup>31</sup>.

The covalent bonding of DNA molecules on surfaces (second approach) alleviates the aforementioned limitations in physical adsorption. A frequently used chemical group is thiols on ultra-flat gold substrates because of the high affinity and stability of S-Au bonds (~ with binding energy of 44 Kcal/mol) over a long period of time<sup>31,36</sup> . However, there are other covalent immobilization and surfaces that are based on the following covalent bonds; carboxyl-amine<sup>37,38,34</sup>, thiol-maleimide<sup>39</sup>, silane modified DNA molecules immobilized on silicon dioxide wafers<sup>40</sup>, amine-aldehyde coated substrate<sup>41-43</sup>, mercaptosilane-thiol bond<sup>44</sup> and isothiocyanate-amine bond<sup>45</sup>. However, the covalent-bonding mechanism through S-Au is widely used in the generation of electrochemical based DNA biosensors<sup>46</sup>.

The last mechanism is based on the streptavidin-biotin complex interaction, which is the subset of non-covalent immobilization mechanism. It is the strongest non-covalent binding between proteins and ligands, whose association constant is of the order of  $10^{15}$  M<sup>32,47,48</sup>. The immobilization procedure involves the following; the surface is first coated with biotin-terminated alkanethiols (if the surface is gold, or biotinylated coated glass slides<sup>49</sup>), followed by the addition of streptavidin and lastly, the introduction of biotin-modified DNA molecules<sup>32</sup>. This mechanism is widely used in generation of DNA microarrays but the production cost is quite expensive because of the multiple steps involved in the synthesis of streptavidin<sup>32</sup>.

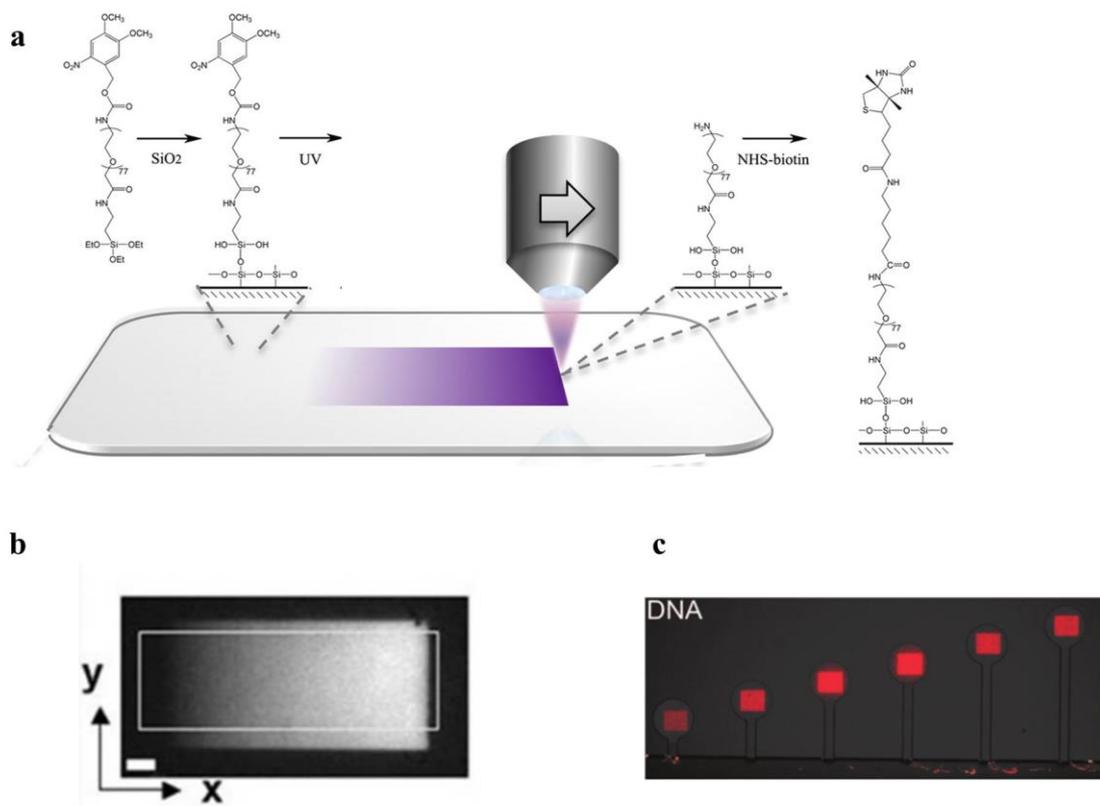
Out of the three mechanisms described above, the S-Au bond formation for immobilizing DNA molecules on surfaces is frequently used to form self-assembled monolayer (SAM) of DNA molecules on an ultra-flat gold substrate<sup>3,36</sup>. SAMs are ordered molecular assemblies that are formed by the spontaneous chemisorption of the molecular constituents from the solution onto a metal (typically gold) surface<sup>48-50</sup>. The basic constituents for DNA molecules to form self-assembled monolayers on the gold surface are one or more sulphur head group (SH) that provides surface bonding along with high lateral mobility; the alkane backbone (e.g. (CH<sub>2</sub>)<sub>6</sub>) that serves as spacer and increases the stability of DNA immobilisation and, lastly, the end group or functional group (COOH, NH<sub>2</sub> etc.), which serves to anchor the DNA to another functional molecule, which modifies the properties of the SAM's topmost interface.

Nuzzo and Allara pioneered SAMs<sup>51</sup> by reporting the monolayer formation of organic di-alkyl disulphide molecules via disulphide dissociation on gold substrates. Subsequently, Bain and Whitesides<sup>52</sup> developed a model system in which linear alkyl thiols were adsorbed from the solution onto the gold to form closely-packed, oriented monolayers. Beyond the feasibility and the formation of ordered self-assembled monolayers of alkanethiols on gold substrate, Scoles and co-workers<sup>53,54</sup> elucidated the underlying mechanism of Au-S interactions and its role in forming highly ordered and closely-packed structure of alkanethiols on Au(111), by using the combination of density functional theory (DFT)-based molecular dynamics simulations and grazing incidence x-ray diffraction (GIXD) techniques. Their findings showed the existence of surface complexes, which involves the joining of two sulphur atoms through a gold adatom (S-Au-S), and interstitial defects such as vacancies at the gold interfacial layer. In addition to the intra-molecule interactions within the alkyl chains, they argued that the appropriate population of vacancies, number of adatom per unit cell, and a limited conversion between RS-Au-SR to

RS bridge structure, are consequential in the formation of superstructure of alkyl thiols on gold substrate. The extension of the linear alkyl thiols model system to nucleic acids was pioneered by Tarlov and coworkers<sup>24</sup> with a study in which they successfully formed self-assembled, ordered monolayers of thiol-derivatised DNA molecules over planar gold substrates. Since then, such a method has become a fundamental approach to immobilize nucleic acid and protein biomolecules on surface towards forming artificial self-assembling nanostructures, and has allowed elucidating important applications in a broad range of research areas, from biosensing<sup>55</sup> to molecular computing<sup>56</sup>.

Another important methodology that has aided the immobilization of biomolecules on surfaces is nanolithography, which, in particular, in the last two decades, has yielded fruitful methods for constructing and studying self-assembled nanostructures of DNA molecules on surfaces<sup>57-59</sup>. Nanolithography belong to the family of top-down fabrication approaches, and involves the manipulation of atoms or molecules on surfaces, using short optical wavelengths (e.g. soft X-rays) or scanning probes to create spatially resolved and electronic addressed patterns on surface with at least one lateral dimension 1-100 nm<sup>60</sup>. The core technology was introduced to generate complex integrated electronic circuits (e.g. modern CPUs) over chips of crystalline silicon<sup>36,61</sup>. Out of many nanolithography methodologies, these three have been frequently used in DNA nanotechnology research field due to their reliability and reproducibility of results; Chemical photolithography, Dip-pen nanolithography and Nanografting. The last two methodologies are scanning probe nanolithography-based methods.

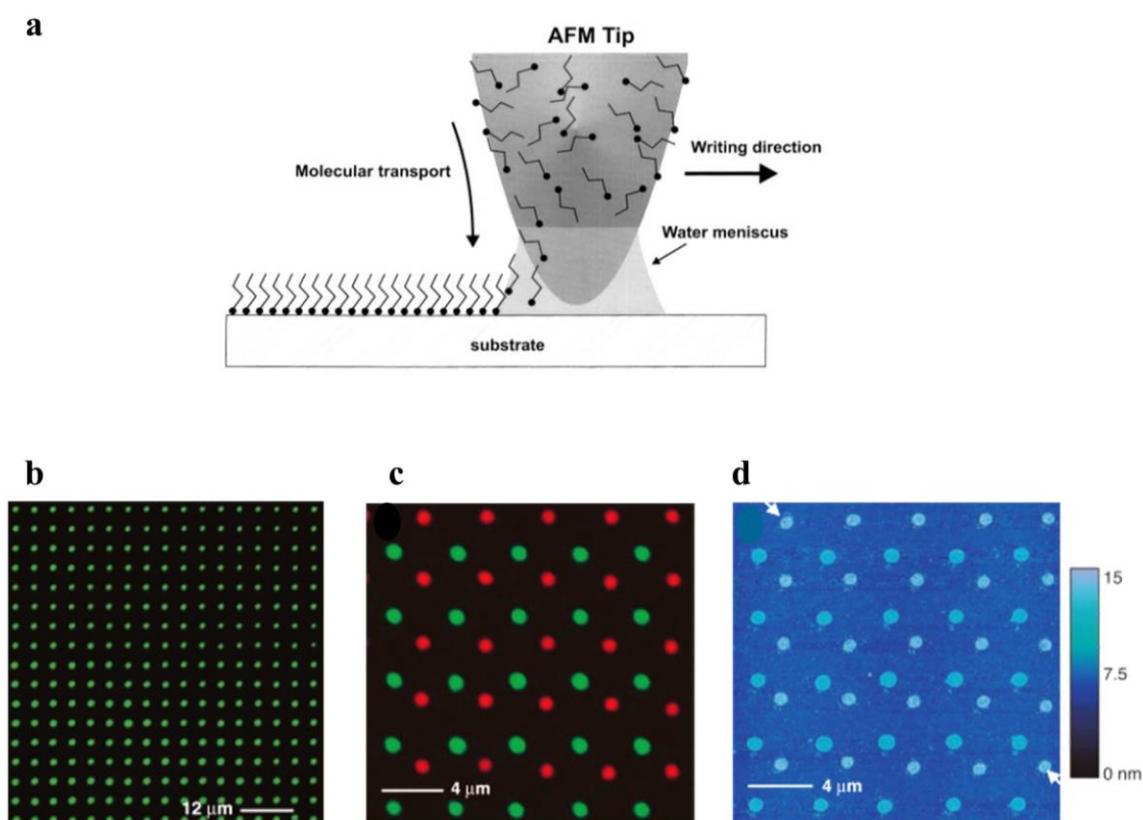
A significant contribution to chemical photolithography came from Roy Bar-ziv and his coworkers<sup>29</sup> in 2007. It involves the assembling of bio-functionalized polyethylene glycol (PEG) on silicon dioxide surface, terminated with photolabile group that protects a reactive amine at one end and an organo-silane as the head group that binds to the surface. Subsequently, the protected amine group is then deprotected using ultra-violet light (of 365 nm wavelength with energy flux of  $\sim 2500$  J/cm<sup>2</sup>), followed by the addition of N-Hydroxysuccinimide (NHS) that helps in linking the amine to biotin, which specifically binds to the streptavidin, followed by biotinylated DNA conjugates (as shown in Fig. 1.2a). By controlling the UV-exposure time and intensity, these procedural steps result into density-dependent, confined, and localised DNA compartment on surface with submicron resolution as shown in Fig. 1.2b. Several shapes<sup>28,29,62</sup> and dimensions<sup>63</sup> of such spatially resolved DNA compartments have been generated and used for different studies that span across gene expression<sup>13,28,63</sup> to cell-free protein biosynthesis<sup>29</sup>.



**Figure 1.2:** Chemical photolithography for the construction of spatially confined DNA compartments<sup>64</sup>. (a) Shows the procedural steps in generating spatially resolved DNA compartment via chemical photolithography, yielding density–gradient DNA compartment as shown in (b) and these DNA compartments can be integrated into microfluidics (see (c)) for gene expression<sup>13</sup> and protein biosynthesis<sup>29</sup>.

Another well-established lithography technique for application to DNA immobilization is termed Dip-pen nanolithography (DPN), and was invented by Chad Mirkin and co-workers<sup>30</sup> in 1999. It uses AFM tips as a pen, a solid substrate (e.g. gold, silicon) as a paper and the chemically modified molecules (e.g. thiol-modified DNA, silane-modified organic molecules) that have high affinity for the substrate as an ink<sup>30,65,66</sup>. This procedure is normally carried out in ambient condition, in which the molecules on the tip are diffusively transported through the water meniscus on to the substrate, mostly driven by capillary forces<sup>65,66</sup>, as shown in Fig. 1.3a. The versatility of DPN to write patterns of different types was first demonstrated with alkanethiols, specifically 1-octadecanethiol (ODT) and 16-mercaptohexadecanoic acid. These molecules were successfully deposited on gold substrate, forming confined self-assembled nanodots on surface with sub-nanometre resolution<sup>65</sup>. Later, the technique was extended to other organic molecules, especially to DNA<sup>67</sup>, peptides<sup>68</sup> and proteins<sup>65,68,69</sup>. Besides gold substrate, other substrates offering

alternative covalent bonding mechanisms can be used for the generation of molecular nanopatterns by means of DPN. For example in Fig. 1.3b-d, a ssDNA molecules modified with an acryamide group is deposited with DPN to covalently bond to thiols on the 3-mercaptopropyltrimethoxysilane (MPTMS) activated SiO<sub>2</sub> substrate. This modification process was done prior to the DPN process<sup>67</sup>.



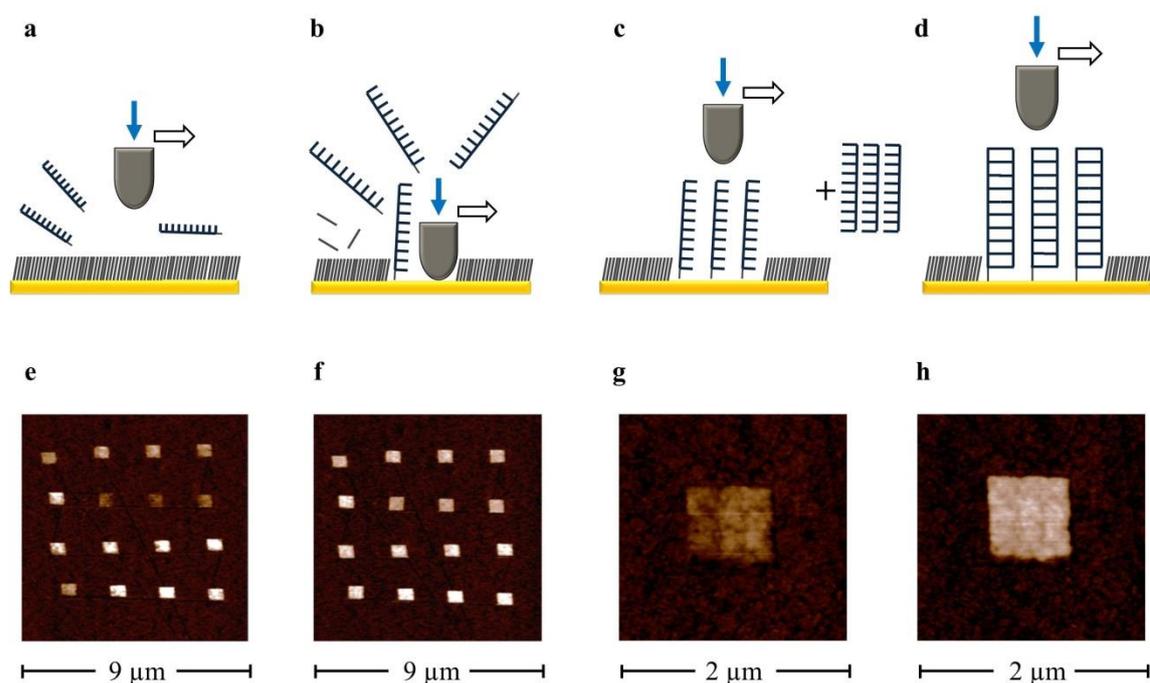
**Figure 1.3:** Dip pen nanolithography. (a) Schematic diagram that shows the mechanism of transporting chemically modified molecules onto the substrate, through the water meniscus, driven by capillary forces, forming self-assembling molecules on the substrate. (b) Shows an epifluorescence micrograph of DNA-nanospot generated via DPN on SiO<sub>2</sub> surface, hybridized with green fluorophore-labeled complementary DNA molecules<sup>67</sup>. (c) Shows a possibility of having multiple DNA ink writing by DPN on the substrate. The epifluorescence micrograph shows an array of two different DNA sequences deposited on SiO<sub>2</sub> surface by DPN and subsequently hybridized with two different (green and red) fluorophore-labelled complementary DNA molecules. (d) AFM micrograph of (c) with assembled gold nanoparticles after the de-hybridization of fluorophore-labeled DNA molecules.

The last scanning probe nanolithography technique that I will present in this section is *Nanografting*. It is in several ways similar to DPN but was pioneered by Gang-yu Liu and her coworkers<sup>31</sup> in 1997, and typically allows three, consecutive and convenient, procedural

steps that cannot be implemented with DPN: 1.) AFM tip scans across an alkanethiol passivated substrate at a low force in the solution that contains another thiol-modified molecules (not necessarily different from the existing one on surface); 2) A higher force (higher set-point voltage, which implies higher energy dissipated through tip-surface friction) is applied within a selected area of the surface, which induces a substitution reaction between the existed alkanethiol in the selected area and the biopolymer in solution, resulting in the assembly of latter over the area scanned by the AFM tip; and 3.) the scan area that includes the location of the nanografted thiol is again scanned at low force to allow that the product of the nanolithographic step is characterized by means of AFM imaging<sup>31,70</sup>. The first demonstration of this technique was with alkanethiols. Liu and her co-workers reported successful nanografting of octadecathiols<sup>31,70</sup>, Docosanethiol<sup>70</sup>, and 16-mercaptohexadecanoic acid<sup>70</sup> within decanethiol monolayer passivated gold surfaces. Their results yield several spatially confined nano-islands with different dimensions<sup>31</sup>, geometries<sup>70</sup> and high edge resolution of ~ 1 nm depending on the radius of curvature of the AFM tip.

In the subsequent years, Gang-yu Liu and co-workers extended the application of nanografting by constructing confined single stranded DNA molecules with different oligonucleotide sequences within the bio-repellent decanethiol self-assembled monolayer on gold substrates<sup>58,71</sup>, using the schematics diagram in Fig. 1.4a-d. They qualitatively studied the functionality, composition of the confined DNA nanostructures and the accessibility of the confined DNA nanostructures to the subsequent biochemical reactions, such as enzymatic digestion (using DNase 1 enzyme)<sup>71</sup> and hybridization<sup>58</sup>, via AFM height measurement carried out before and after each biochemical reactions. Also, they showed the possibility of constructing density-dependent confined DNA nanostructures by fine-tuning the following parameters; 1) concentration of the ssDNA solution introduced on to the surface before the nanografting process; 2) the applied force used during the nanografting process; and, lastly, the number of times the AFM over-write a selected area during the nanografting process<sup>58</sup>. Their qualitative results are summarized as follows; 1) the composition of confined DNA nanostructure is purely DNA molecules as the enzyme fully digested the confined DNA nanosystem and that the DNA molecules are accessible for the further reactions<sup>71</sup>; and 2) the ssDNA molecules within the confined nanostructures maintain their bioactivity, selectivity and specificity towards the complementary ssDNA during hybridization steps, however, the hybridization efficiency is highly dependent on the parking density of the confined DNA nanostructures<sup>58</sup>.

Beyond the pioneering work by Gang-yu Liu and her co-workers, nanografting is now an unparalleled nanolithography tools for construction of spatially resolved DNA nanostructures<sup>20,27</sup> and protein arrays<sup>72-74</sup> that are used for quantitative characterization of biochemical reactions.



**Figure 1.4:** Construction of confined DNA compartment via Nanografting. The schematic diagram shows that in (a) AFM tip scan (at low force) across the ethylene glycol-terminated alkyl thiol monolayer assembled on an ultra-flat gold substrate within the solution of thiol-modified DNA molecules with hexane-linker. Next, at high force (high set –point voltage), the AFM tip plows a selected section of the monolayer and grafted instead the thiol-modified ssDNA from the solution, resulting into confinement of DNA molecules in the mixed of ethylene glycol-terminated alkyl thiol monolayer, as shown in (c), where the AFM tip scan at a low force the scan area that contains the confined DNA nanostructure. Lastly, a subsequent reaction can be done by introducing complementary ssDNA molecules that can result into the formation of hybridized DNA molecules confined on surface. (e) Shows the AFM micrograph of 4 X 4 arrays of confined ssDNA nanostructure, with different density generated by nanografting, followed by the AFM micrograph of the (e) after hybridization process. (g) Shows the AFM micrograph of one of the confined DNA nanostructure in (e) and the resulting nanostructure of the same after hybridization in (h). The AFM micrographs from 1.4e-h are experimental results obtained by me when working on the DNA-DNA affinity within confined and density dependent DNA arrays in our laboratory.

As anticipated, nanografting is an alternative technique to Dip pen nanolithography, the former transfers energy onto the alkanethiol-passivated substrate, while the latter

exploits transportation of molecules on surface without prior passivation<sup>65</sup>. Here are some of the advantages of nanografting over DPN ; 1) based on the energy transfer associated with nanografting, it facilitates the kinetic and spontaneous organization of adsorption of molecules, by lowering the activation energy for spatially confined self-assembling process; 2) it gives the possibility for identifying the molecular orientation by measuring the height of the confined nanostructure with nanometre precision; 3) it provides the tools for studying the functionality and the biophysical properties of the confined biopolymer by constructing density-dependent confined nanostructures at addressable loci on the same surface; 4) lastly, it allows label-free detection of molecular recognition or interactions within the nanostructures by means of AFM imaging. In addition to the advantages of nanografting over DPN, both techniques have an edge over optical nanolithography, as the latter is mainly limited in resolution by the diffraction limit of the optics, while this does not apply to the former techniques. In addition, the spatial resolution of dip-pen nanolithography significantly depends on the relative humidity, which directly affects the width of the water meniscus between the writing tip and the substrate, thus influencing the rate of deposition of the molecules onto the substrate<sup>66</sup>.

### **1.3 Structure—function relationship in surface-bound DNA nanostructures; Effect of confinement, crowding, and ionic strength on the DNA conformation and DNA hybridization**

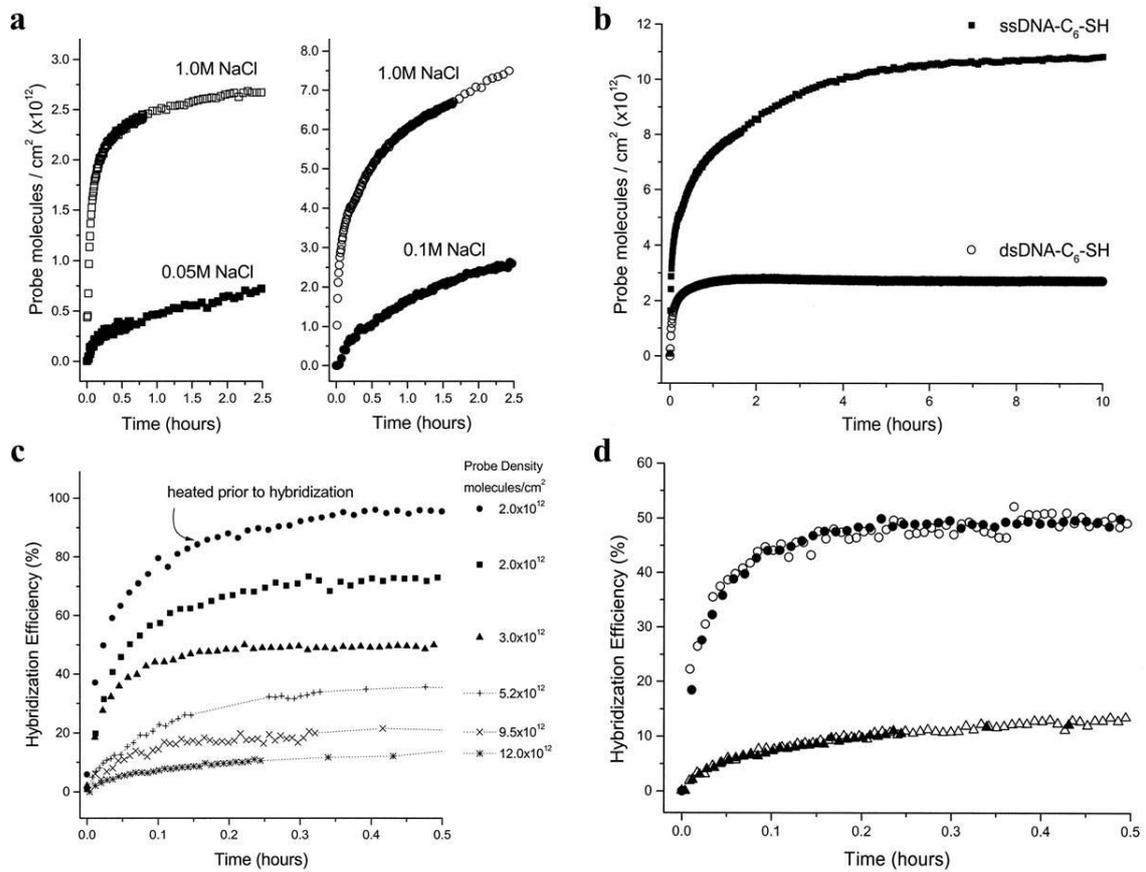
Here is a review of selected literatures, which have used the aforementioned mechanisms for immobilization of DNA molecules on surfaces. Also, we highlight their findings as they reveal the effect of crowding, ionic strength and confinement on the functionality of surface-bound DNA nanostructures. Moreover, it is very important to understand the physico-chemical and structural properties associated with surface-bound DNA nanostructures, as these will aid the development and improve the effectiveness of such nanostructures for frontier application in synthetic biology and nano-biotechnology.

Starting with the pioneer work by Tarlov and coworkers<sup>24</sup>, in which the thiol-derivatized DNA molecules were spontaneously chemisorbed onto the gold substrate, to form self-assembled monolayers. They examined the effect of the coverage density on hybridization of DNA molecules on surfaces, by forming the mixed monolayer of DNA/ 6-mercapto-1-hexanol (MCH), where MCH serves as spacer within the monolayer DNA/MCH monolayer. With such mixed monolayer configuration, they characterized the monolayer using X-ray photoelectron spectroscopy (XPS), ellipsometry, and P-

radiolabelling technique. Their findings showed that hybridization process could not take place in the highly dense full (un-mixed) ssDNA monolayer because of the densely-packed configuration of the monolayer, due to strong electrostatic screening inhibiting the achievement of physical proximity, and so preventing the chemical recognition between the complementary strands. However, they showed the hybridization efficiency to be greatly enhanced in the mixed DNA/MCH monolayer, because the presence of the spacer (MCH) able to increase the accessibility of the complementary strands in the monolayer. Also, they observed an influence of salt concentration on the surface coverage in the DNA monolayer, meaning the higher the salt concentration, the more the density of the probes immobilized on the gold substrate. Gamble and coworkers<sup>75</sup> further buttress the role of MCH in the mixed DNA/MCH monolayers on gold substrate. Using the same techniques employed by Tarlov<sup>24</sup>, and in combination with near-edge X-ray absorption fine structure spectroscopy (NEXAFS) and fluorescence intensity measurements, their finding showed that the back filling of MCH molecules fill up the surrounding vacant space within the DNA monolayer. Also, the presence of MCH diluents enhanced the upright orientation of the single stranded DNA molecules, therefore improving the hybridization process in such DNA/MCH mixed monolayer.

Still elucidating the effect of surface density on hybridization and the dependency of probe density on the ionic strength of the DNA solution, Peterson<sup>76</sup> *et. al.* employed Surface Plasmon Resonance (SPR) spectroscopy to monitor the kinetics associated with the immobilization of ssDNA on surface and as well as the subsequent hybridization process. Using the same immobilization mechanism, namely of thiol-modified-ssDNA on gold substrates, they varied the probe density using the following procedures; 1) by varying the incubation time of the gold substrate in the solution of thiol-modified ssDNA; 2) by varying the ionic strength of the solution of single stranded DNA and 3) by applying an attractive electrostatic field at the interface to assist in the immobilization of negatively charged single strand DNA. Their findings are summarized as it follows; 1) at high ionic strength (1 M NaCl), for both dsDNA and ssDNA molecules, they observed higher probe density, and this is because of the presence of large monovalent ions that help in screening the electrostatic repulsion between the molecules (see Fig. 1.5 a). However, at low ionic strength (0.05-0.1 M NaCl), they obtained low probe density because of the high electrostatic repulsion between the DNA molecules. This is in great agreement with the findings of Tarlov<sup>24</sup> *et. al* ; 2) With respect to the effect of incubation time, they observed an exponential increase of probe density of ssDNA molecules on the surface as a function of time, while the probe

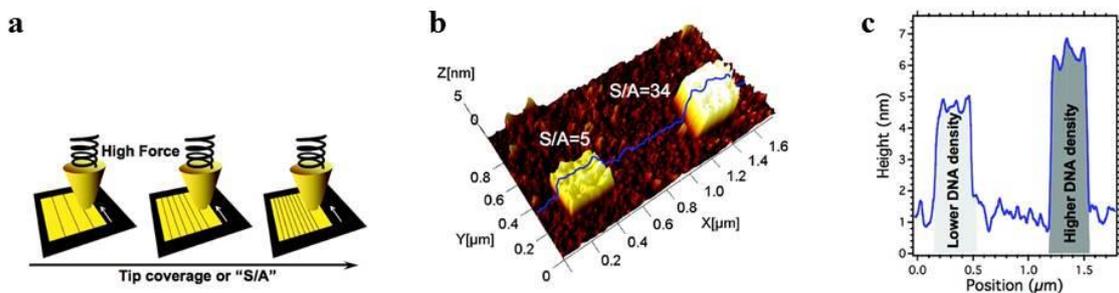
density of dsDNA remains constant over the same period of time with lower probe density compared to ssDNA (see Fig. 1.5 b). Lastly, they reported that surface probe density is a determinant factor for successful hybridization process: in the low probe density regime, they observed a faster kinetics leading to ~100% hybridization efficiency with the complementary strands in solution, while at high density regime, a much slower kinetics limited the hybridization to ~10% (see Fig. 1.5 d).



**Figure 1.5:** The effect of probe density on hybridization. The plot (a) shows the comparison of probe immobilization kinetics as a function of ionic strength formed from solutions containing  $1 \mu\text{M}$  dsDNA-C6-SH (on the left of (a)) and  $1 \mu\text{M}$  ssDNA-C6-SH, which is the plot on the right of (a). In (b), the two curves represent the immobilization kinetics of ssDNA-C6-SH (closed squares) and dsDNA-C6-SH (open circles) from  $1 \mu\text{M}$  DNA solutions in  $1 \text{ M KH}_2\text{PO}_4$ , showing the comparative plot of the probe density of dsDNA and ssDNA as a function of time. (c) The plot shows the dependence of hybridization efficiency on the probe density of ssDNA and the influence of heating prior to the hybridization process. Finally, (d) shows the comparison of the hybridization kinetics for the probe density formed by different immobilization strategies. Hybridization kinetics are shown for two low probe density films,  $\sim 3 \times 10^{12}$  molecules/ $\text{cm}^2$  (circles) and two high probe density films,  $\sim 1.2 \times 10^{13}$  molecules/ $\text{cm}^2$  (triangles). Probe films at low density, formed either from ssDNA-C6-SH immobilization (filled circles) or via dsDNA-C6-SH immobilization (open circles) show

comparable kinetics and overall efficiency of capture. Similarly, probe films formed via immobilization of ssDNA-C6-SH under electrostatic conditions of +0.3 V (open triangles) or via incubation overnight at open circuit (filled triangles) are indistinguishable. All hybridization kinetics obtained under the same conditions of 1  $\mu$ M target in 1MNaCl with TE buffer. Adapted from reference <sup>76</sup>. Copyright © 2001 by Oxford University Press.

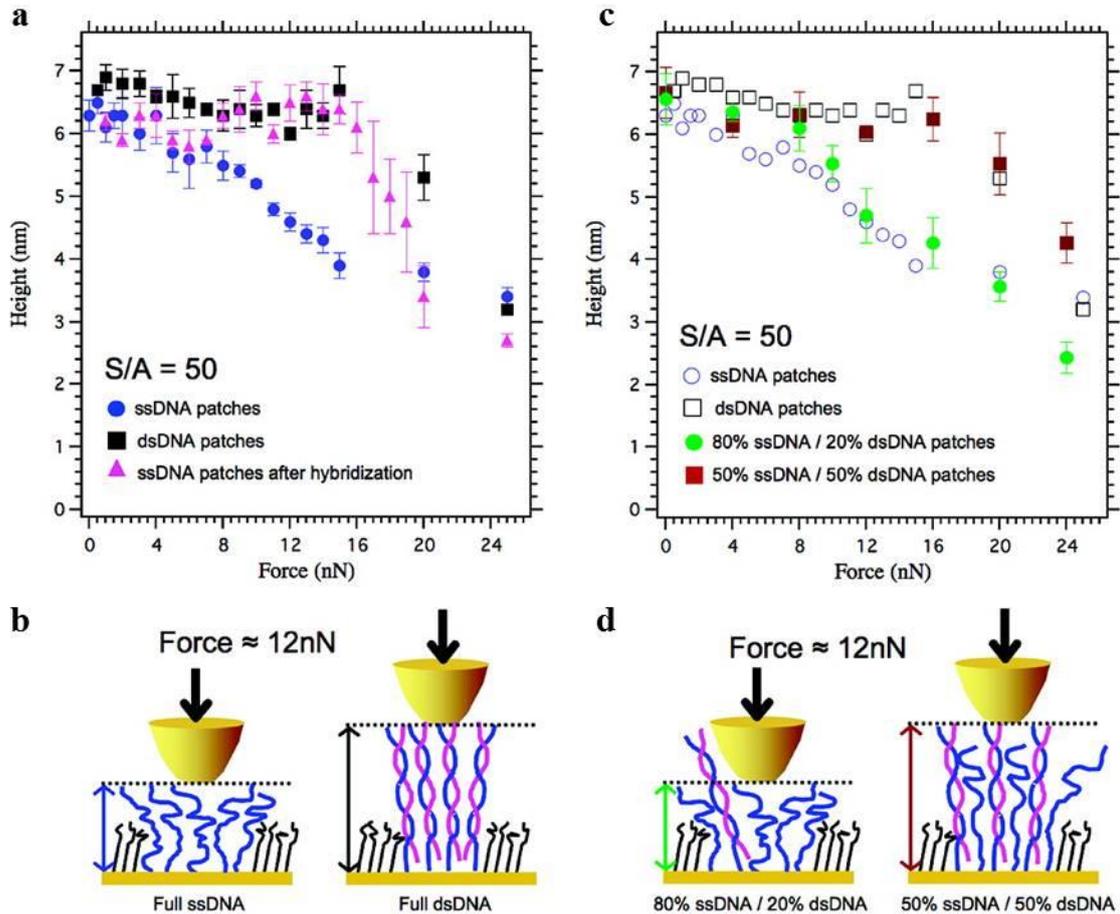
On the other hand, Scoles and co-workers studied the effect of surface coverage on the functionality of the spatially confined DNA nanostructures, generated by AFM Nanografting<sup>20</sup>. In this study, spatially confined ssDNA nanostructures were created within monolayer of bio-repellent oligo-ethylene glycol on an ultra-flat gold surface, via the procedural steps of nanografting described in the previous section. Building on the nanografting procedures developed by Liu<sup>70,71</sup>, they were able to fine-tune the surface density of the spatially confined DNA nanostructures by varying the number of times that the AFM tip over-write a given area during the high-force scanning process (see Fig. 1.6 a). The investigators defined the surface density parameter as  $S/A$ , where  $S$  stand for the scanned area of the AFM tip during the nanografting and  $A$  represent the final actual area of the confined nanostructure. This parameter  $S/A$  was calculated by considering the radius of curvature of the tip in contact with the surface, the number of pixels or scanned lines during the nanografting process and the length or width of the generated confined nanostructures. With such surface density calibration, they vary the  $S/A$  from 1-60 and by measuring the height of the confined ssDNA nanostructures using the side-by side AFM height measurement, their findings show that the height of the confined ssDNA molecules varies proportionally to the  $S/A$ , meaning, the higher the number of times the AFM tip over-write a given area during nanografting, the more the number of DNA molecules grafted within the area and the higher the density of the obtained DNA nanostructures (see Fig. 1.6 a-c).



**Figure 1.6:** Fine-tuning the packing density of confined DNA nanostructures during nanografting process. (a) Illustrate the methodological approach employed to create highly dense confined nanostructures by varying the number of times the AFM tips over-write a selected area during nanografting process. An example of two nanografted ssDNA patches are shown in (b), one at low

density  $S/A = 5$  and the second at high density with  $S/A = 34$  and the equivalent heights as measured using side by side AFM height measurement are shown in (c). Adapted from reference <sup>20</sup>. Copyright © 2008 American Chemical Society.

Following their approach, the authors investigated the density-dependent reactivity of the laterally confined ssDNA nanostructures upon hybridization, involving complementary ssDNA molecules of the same length (18 nt). At low  $S/A$  (low density), the height of the nanostructures increase after the hybridization process, signifying the stiffness increase of the dsDNA as to compare to the ssDNA. However, at high  $S/A$  (high density), where the height of the nanopatches have already reach saturation, there was no change in height. In order to know whether hybridization takes place or not at such high density, they employed compressibility measurements, in which the height measurements were taken at appreciable different imaging force. Their results show that the mechanical response of confined nanografted ssDNA under varying imaging force is totally different to the mechanical resistance of dsDNA of the same  $S/A$  (see Fig. 1.7 a-b). In addition, after hybridization, the mechanical resistance of the formerly ssDNA nanografted patches increased to unequivocally match the mechanical resistance of dsDNA nanografted patches as shown by the height vs. applied imaging load in Fig. 1.7 a-b. Lastly, by comparing the mechanical resistance of the nanopatches that consists of the predefined mixed composition of the ssDNA and dsDNA molecules (see Fig. 1.7 c-d), the authors show that the hybridization efficiency in the highly dense nanografted ssDNA is at least 50%, as shown in the Fig. 1.7 c. In partial disagreement with the conclusions of Peterson<sup>76</sup> *et. al.* study in 2001, this study shows that the hybridization efficiency in high density ssDNA nanografted patches can be much higher than the value measured in self-assembled ssDNA monolayer formed spontaneously (only ~10 %).



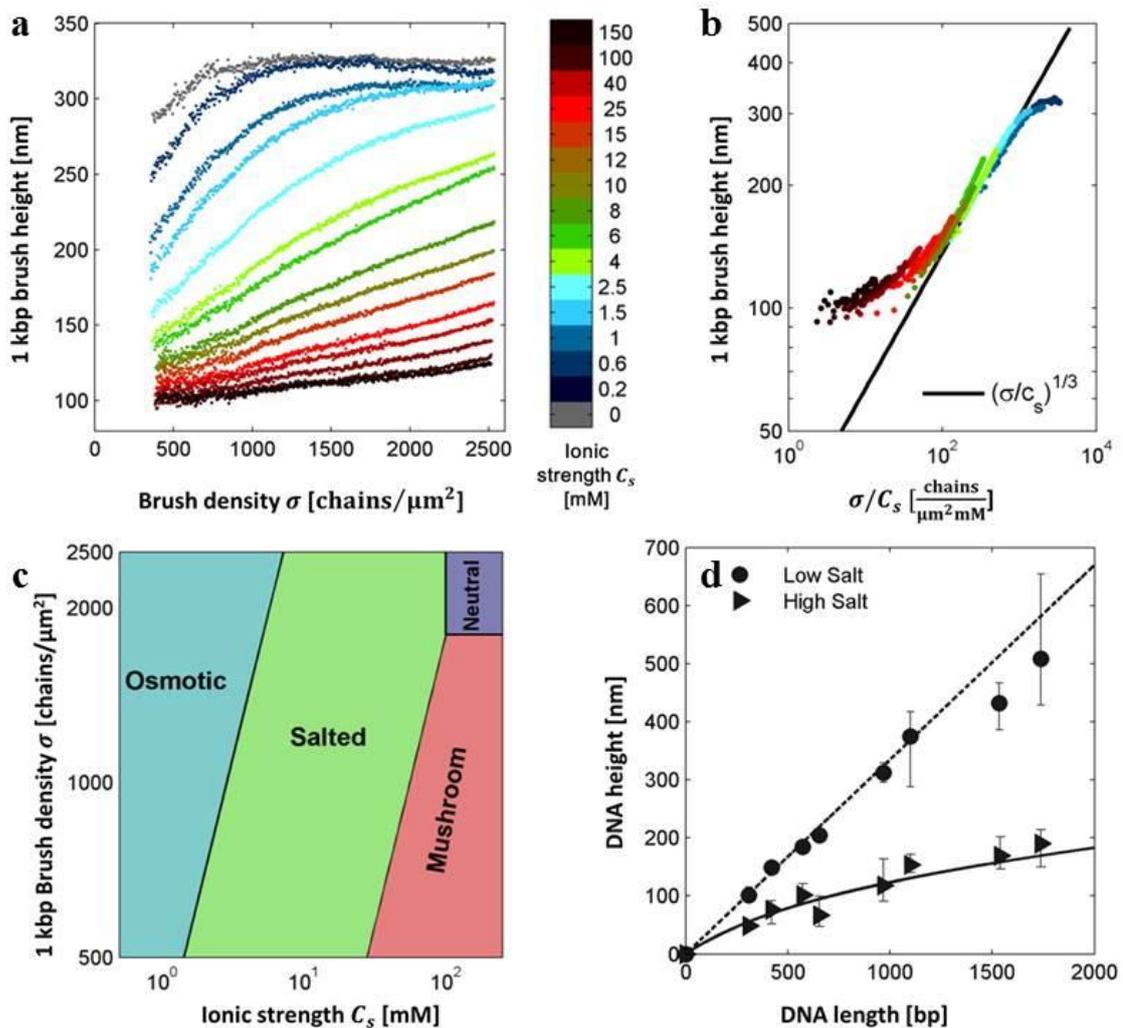
**Figure 1.7:** Mechanical responses of highly dense confined DNA nanostructures before and after *in situ* hybridization process. (a) Shows the compressibility curves, generated by AFM topographic height imaging of highly dense DNA nanopatches as a function of applied imaging force. The blue-coloured circles represent the mechanical resistance of the ssDNA nanopatch before the hybridization while the pink-colored triangles represents the mechanical resistance of the same ssDNA patches after hybridization process. Obviously, the mechanical resistance of the ssDNA nanopatches consistently matches the mechanical resistance of the dsDNA nanopatches of the same surface density. (b) The schematic depicts the mechanical response of the ssDNA nanopatches and the dsDNA nanopatches under an AFM imaging force of 12 nN. (c) Shows the compressibility curves of nanografted patches with mixed composition i.e 80% ssDNA: 20% dsDNA and 50% ssDNA: 50% dsDNA. These curves are compared with the full ssDNA and full dsDNA nanopatches, showing the overlapping behaviour of mechanical resistance from full dsDNA nanopatches (open black squares) and mixed 50% ssDNA/50% dsDNA nanopatches (filled brown squares). The schematics in (d) depicts the mechanical response of nanografted patches with mixed composition, 80% ssDNA: 20% dsDNA and 50% ssDNA: 50% dsDNA, under the same imaging force as those in (b). Comparing (b) and (d), the mechanical response of full dsDNA and mixed 50% ssDNA/50%

dsDNA nanopatches are nearly the same. Adapted from reference <sup>20</sup>. Copyright © 2008 American Chemical Society.

A recent work by Nkoua Ngavouka <sup>26</sup> and coworkers examined the role of cation-electrostatic screening and molecular crowding on the conformational stability of spatially confined single stranded (ss) DNA molecules (44 bps long), generated through nanografting. Using the nanografting parameters established by Liu<sup>58</sup> and Scoles<sup>20</sup>, the authors constructed different densities of laterally confined ssDNA within the monolayer of bio-repellent ethylene glycol terminated alkyl thiol on the gold surface. The density range considered in this study spans from  $1 \times 10^{12}$  to  $2 \times 10^{13}$  molecules/cm<sup>2</sup>, and they then monitored the changes in topographic height of the density-dependent nanografted ssDNA patches upon changing the salt type from monovalent to divalent cation and the concentration of these ionic solutions, via the label free AFM height measurements. Their findings show that highly dense nanografted ssDNA patches, with density  $2 \times 10^{13}$  molecules/cm<sup>2</sup>, have a stretched conformation in vertical direction in low ionic strength (NaCl < 100 mM, analogous to Debye length > 1 nm) solution, with maximum height of ~ 15 nm. However, the reverse is the case in the high ionic solution (NaCl > 100 mM, corresponding to Debye length < 1 nm), with the same high density, the tethered ssDNA molecules within the highly dense confined DNA molecules have mushroom-like (compact) conformation with a reduced height (~ 5 nm), because of the effective screening of the backbone charges on the ssDNA by the monovalent ions in solution. Furthermore, they fit their experimental data with simplified Daoud–Cotton model (DCM) that was developed by Hariharan *et. al.* for spherical brushes, and estimated a hypothetical value for the average inter-nucleotides distances for such a confined ssDNA nanopatches, to be  $d = 0.6 \pm 0.03$  nm, specifically in NaCl-containing solution. Also, they changed the salt in solution, from monovalent to divalent (CaCl<sub>2</sub>, MgCl<sub>2</sub>), and measured the topographic height of densely packed ssDNA nanopatches. They observed a notable difference in height of the confined ssDNA nanopatches with a direct reduction in the inter-nucleotide distance ( $d = 0.4 \pm 0.02$  nm in CaCl<sub>2</sub> and  $d = 0.36 \pm 0.02$  nm in MgCl<sub>2</sub>), showing that divalent counterions are more effectual in screening the negative surface charges on the backbone of DNA molecules.

Bar-ziv<sup>64</sup> *et al.* also investigated the effect of cation-counterions on the physics and the conformation of the spatially resolved DNA brushes, constructed on silicon-dioxide substrate using chemical photolithography<sup>29</sup>. Using the approaches described under

chemical photolithography in the previous section, they formed densely packed rectangular patterns of  $\sim 1$  kbps dsDNA of dimension  $100 \times 50 \mu\text{m}^2$  with packing density ranging from 400 to 2500 chains /  $\mu\text{m}^2$ . They monitored the height of DNA brushes, with varying surface density, in different concentration of NaCl solution, via total internal reflection spectroscopy. Their findings show that the DNA brush has a stretch conformation in water (the absence of salt, see Fig. 1.8 a) with height  $\sim 340$  nm, independent of the surface density of the DNA brush. Also, in the mid-range of salt concentration,  $0.2 \leq C_s < 100$  mM, the brush began to stretch as the surface density increases. Furthermore, they also measured the variation in height of the DNA brushes as a function of increasing brush-density-to-ionic-strength ratio and observed four distinct conformation states associated with DNA brushes; osmotic pressure regime, salted regime, mushroom regime and quasi-neutral regime (see Fig. 1.8 c). Their possible interpretation for such observed phases is as follows; 1) In osmotic pressure regime, which corresponds to  $C_s < 1$  mM and the brush-density-to-ionic-strength ratio  $> 1000$  chains  $\mu\text{m}^{-2}\text{mM}^{-1}$ , they argued that osmotic pressure due to the confined counterions within the DNA brush in this regime is sufficient to induce maximum extension of the DNA brush, resulting into linear correlation between the height of the DNA brush and the contour length of the DNA as shown in Fig. 1.8 d; 2) In the salted regime ( $100 < \sigma/C_s < 1000$  chains  $\mu\text{m}^{-2}\text{mM}^{-1}$ ), they characterized the behaviour of the DNA brush with respect to ionic strength with a power law, with scaling factor of  $1/3$ . Here, they explained that osmotic pressure counter-balance the entropy of the DNA polymer, which results into no stretching with linear relationship between the height of DNA brush and brush-density-to-ionic-strength ratio (see Fig. 1.8 b); 3) The mushroom regime ( $\sigma/C_s < 100$  chains  $\mu\text{m}^{-2}\text{mM}^{-1}$ ), which is evident in the region of low brush density and high ionic strength, for which they explained that osmotic pressure is insignificant since the brush conforms to a compact state with minimum height, which they characterized with the wormlike-chain model, as depicted in Fig.1.8d; 4) Lastly, in the quasi-neutral regime ( $C_s > 100$  mM,  $\sigma > 2000$  chains/ $\mu\text{m}^2$ ) that corresponds to high ionic strength and high brush density, they observed that the measured height of the DNA brush and the brush density follow a scaling function,  $h_N \propto Npl\rho^{1/3}$ , which they believe should be due to the crowding or excluded volume effect.



**Figure 1.8:** Effect of ionic strength on the conformation and height of spatially confined dsDNA brushes. (a) Comparison of height measurements of DNA brushes as a function NaCl ionic strength (from 0-150 mM concentration) shown in the colour scale, and the surface density. (b) Shows the correlation between the height of the DNA brushes and the surface density-to-ionic strength ratio. The experimental data was fitted with power law ( $h \propto (\sigma / C_s)^\alpha$ ) with scaling factor of 1/3. (c) Represent a phase diagram of DNA brush showing four phases, from osmotic, salted, and mushroom to neutral regime, as a function of ionic strength and surface density. (d) Shows the dependency of the height of DNA brushes on the DNA contour length in relation to the ionic strength. Adapted from reference <sup>64</sup>. Copyright © 2014 American Chemical Society.

The study by Bar-ziv<sup>64</sup> *et al.*, is a complementary work to Nkoua Ngavouka, since the later focus on the conformation of spatially confined ssDNA (44 bps) brushes, while the former focus on the confined dsDNA (~1 kbps) brushes. Nkoua Ngavouka *et al.* only identified two phase regimes for ssDNA brushes, osmotic and salted regime, while Bar-viz *et al.* identified four regimes, even though the brush density range considered by former is

about 2-fold greater than those considered by Bar-ziv. However, these studies have shown that ionic strength and surface density have a significant effect on the conformation and should be considered while designing both ssDNA and dsDNA nano/micro nanostructures.

#### 1.4 Synthetic DNA compartments for biochemical reactions

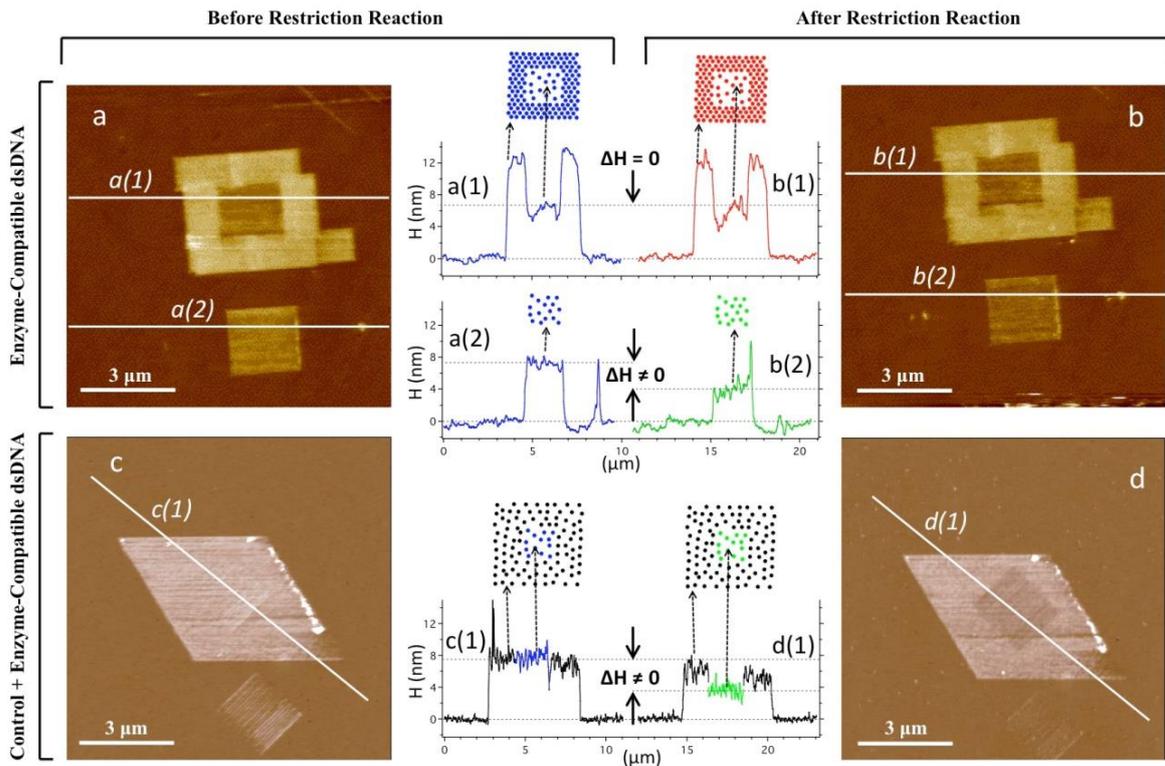
In nature, biomolecules are wired to function in confined, crowded and localized compartments within living organisms<sup>10,77,78</sup>. Conversely, most of the biomolecules are studied in dilute solutions, a state of high degree of freedom without any appreciable effect of confinement and crowding<sup>77</sup>. Another trending approach is by introducing crowding agents to the buffer solution<sup>79</sup>, however this approach is limited by the solubility limit of the crowding agents<sup>80</sup>, and also random diffusion of the crowders within the bulk solution<sup>80,81</sup>, which results into little or no biological information on the effect of nano-confinement and localized compartment on the biochemical reaction.

With respect to the review in the previous section, notably, synthetic spatially resolved DNA nanostructures can be designed and created, with intrinsic conformation and molecular densities, which mimics the crowding, packaging, and compartmentalization of DNA molecules both in prokaryotic<sup>10</sup> and eukaryotic cells<sup>64</sup>. These offer some exciting possibilities to study biomolecules in a synthetic milieu that is akin to the intracellular organization and circuitry and yield biomolecular information that are not captured in standard *in-vitro* assays.

Our group investigated the effect of nanoscale confinement and crowding on the function of restriction enzyme *Dpn II*--an enzyme that cleaves dsDNA molecules specifically at a restriction site 5'-GATC-3'--within spatially resolved DNA nanopatches, that were constructed using AFM nanolithography-based approach, termed nanografting<sup>21</sup>. The investigators employ AFM topographic height measurement to quantitatively detect the digestion of the dsDNA molecules before and after the action of *DpnII* within the confined DNA nanopatches. They observed the effective function of *DpnII* within the relatively low density, characterized with the 50 % decrease in the height of the dsDNA nanopatches after enzymatic reaction. This is in accordance with the position of the restriction sites on the dsDNA molecules, which are the constituents of the confined DNA nanopatches. Also, their findings reveal that the function of *DpnII* is sterically hindered, in a step-wise fashion, when the density of the nanopatch is above a critical density threshold. Based on their interpretation, the inhibition on the *DpnII* action is due to the reduced intra-distance spacing

between the dsDNA molecules at high density, and as such the intra-distance spacing is much less to the dimension of the *DpnII* dimer.

In the subsequent years, the authors further elucidate the mechanism of steric hindrance within the highly dense DNA nanopatches and its direct effect on the diffusion of *DpnII* into the laterally confined dsDNA nanopatches, using AFM topographic height measurements combined with fluorescence imaging<sup>22</sup>. They constructed a composite-like, confined DNA nanostructure, which consists of low-density DNA nanopatch at the centre and fenced with high-density DNA nanopatch, with cognate restriction site for *DpnII*, located in the middle along each dsDNA molecules. With such innovative construct, the authors investigated the mechanism through which the enzyme gains access into the dsDNA nanopatch. Their findings show that enzyme cannot gain access into the cleavable low-density component of the composite-like nanopatch but only access a cleavable low-density nanopatch if the barrier is of low density or absent (see Fig. 1.9). With such result, they concluded that the entrance path of *DpnII* into the confined DNA nanopatch is through the sidewalls, which is few nanometre in size, and not through the topmost interface section, which is far bigger. In addition, they argued that once the enzyme enters into the nanopatch, it diffuses through the nanopatch in x and y direction in a two-dimensional way.

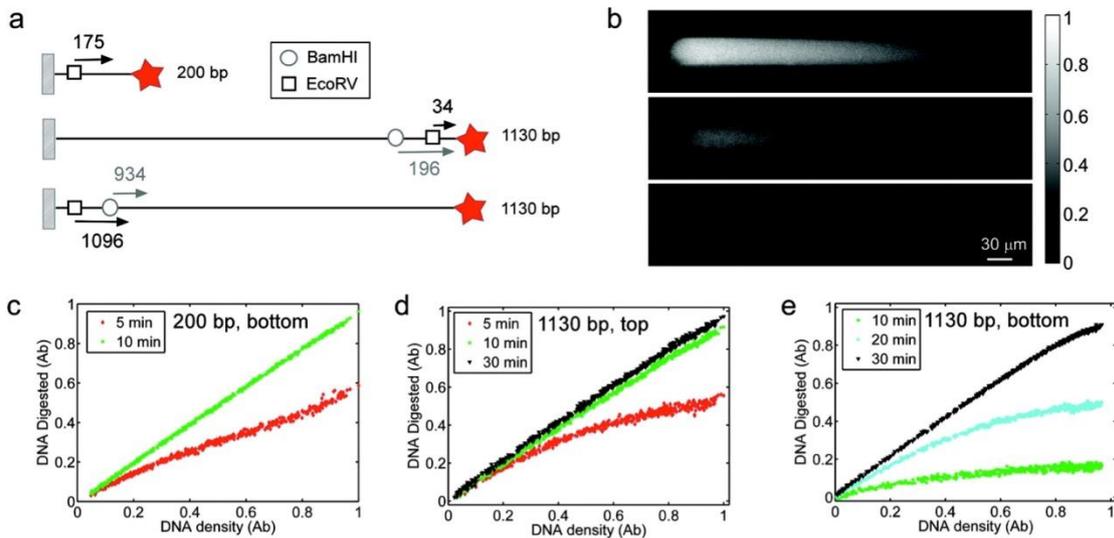


**Figure 1.9:** Detection of two-dimension enzyme diffusion through confined dsDNA nanopatches. In (a) and (b) are the two AFM micrographs that show two distinct nanografted patches; a1 is a

composite nanopatch that contain of low density nanopatch surrounded with high density nanopatch.; a2 is a cleavable low density patch without the high-density component. Likewise, the same AFM micrograph in (a) is shown in b after the DpnII enzymatic reaction. In between a and b are the height profiles that corresponds to the line profile across a1, a2, b1, and b2, together with schematics that depict the density of dsDNA molecules within the composite a1 (blue dots), cleavable low density (blue dots without the fence), composite b1 (red dots), and cleavable b2 (green dots without the fence). Based on the height profile (a1) and (b1), the DpnII could not gain access to the low-density component in composite a1, and as such there was no change in height after enzymatic reaction. However, DpnII gain access into the cleavable low-density nanopatch a2 that is without high density fence, it is obvious in the nanopatch b2 after the enzymatic reaction, as the nanopatch appear darker compare to the same nanopatch before enzymatic reaction. The generated nanostructures in (a) are made up of the same 44 bps long dsDNA molecules, which consists of cognate restriction site for DpnII, located in the middle. In (c) and (d) are two different confined dsDNA nanostructures; The upper of (c) contains the cleavable low density nanopatch surrounded with larger low-density non- cognate nanopatch (made up of non-cleavable dsDNA molecules) while the lower part of (c) contains cleavable low density patch without any surrounding nanopatch. The same constructs in (c) is shown in (d) after the Dpn II enzymatic reaction. In between (c) and (d) are the height profiles across c1 and d1, together with the schematics that represents the density of dsDNA in c1 and d1. Adapted from reference <sup>22</sup>. Copyright © 2011, Rights Managed by Nature Publishing Group.

Bar-ziv<sup>16</sup> *et al.* investigated the digestion efficiency of two distinct type II endonucleases, EcoRV and BamHI, within spatially resolved DNA brushes with varying densities. These DNA brushes are made up of fluorophore-labelled DNA molecules with changing positions of restriction sites (at the bottom or at the top of the molecules) and length of the DNA molecules (200 bp and 1.13kbp) (see Fig. 1.10 a). They utilized chemical nanolithography<sup>29</sup> to construct a continuously-extended DNA surface of varying density, from the maximum density to the minimum density, with intra-molecular distance of 20-30 nm. Using fluorescence microscopy, they monitored the real-time activity of EcoRV and BamHI within DNA brushes with aforementioned design and properties. Their findings show that the activity of endonucleases was density-independent, with no inhibition to the accessibility of the enzymes to reach its recognition site, located at the bottom of the brush in 200-bp DNA brushes (see Fig. 1.10 c). Conversely, in the 1130-bp brush, the position of the recognition site and the density of the brush have a profound effect on the digestion of the enzymes (see Fig. 1.10 d, e); For 1130-bp brushes with restriction site at the top, the full

digestion of the brush was completed in 10 minutes (Fig. 1.10 d), while in the 1130-bp brushes with the restriction site at the bottom, the full digestion was observed at 30 minutes.



**Figure 1.10:** Real time digestion of endonucleases within density-gradient DNA brush. The design of DNA molecules used for the experiment is shown in (a). The 200-bp DNA consists of the restriction site at the position close to the surface (bottom), and fluorophore tag at the other end. The second molecule, which 1130-bp long has two separate designs; one design with restriction site at the top and the second design with restriction site at the bottom. Also, both design for 1130 bp have fluorophore-tag at the other end of the molecule. (b) Shows the real-time fluorescence micrographs of the activity of EcoRV within continuous-density gradient 1130-bp DNA brush with restriction site at the top, before the enzymatic reaction (top micrograph), after 5 minutes of the reaction (middle micrograph) and after 10 minutes of the reaction (last micrograph). (c-e) Shows the plots that represent the time-dependence activity of EcoRV enzyme with different DNA brushes with respect to the fraction of the cleavages and density; (c) 200-bp DNA brush with the restriction site at the bottom, (d) 1.13-kbp DNA brush with restriction site at the top and lastly (d) 1.13-kbp DNA brush with restriction site at the bottom. Copyright © 2009 American Chemical Society.

Comparing the results obtained by our group and that of Bar-ziv, the difference could be due to the different mode of methodologies for the construction of the confined DNA brushes and the detection of the enzymatic activity. Also, the density range considered by our group is far higher than those in Bar-ziv work.

With respect to the reviews, they all studied DNA-binding proteins in cognate spatially resolved DNA brushes, in this thesis, we take a step further by harnessing the inherent properties of DNA compartment; confinement, heterogeneity and homogeneity molecular density, to fine-tune the specificity and the recognition of site-specific DNA-

binding protein. This contribution will expand the frontier of synthetic DNA brushes in biotechnological applications.

### **1.5 Translation of DNA arrays into peptide chips for biomarker detection and protein recognition**

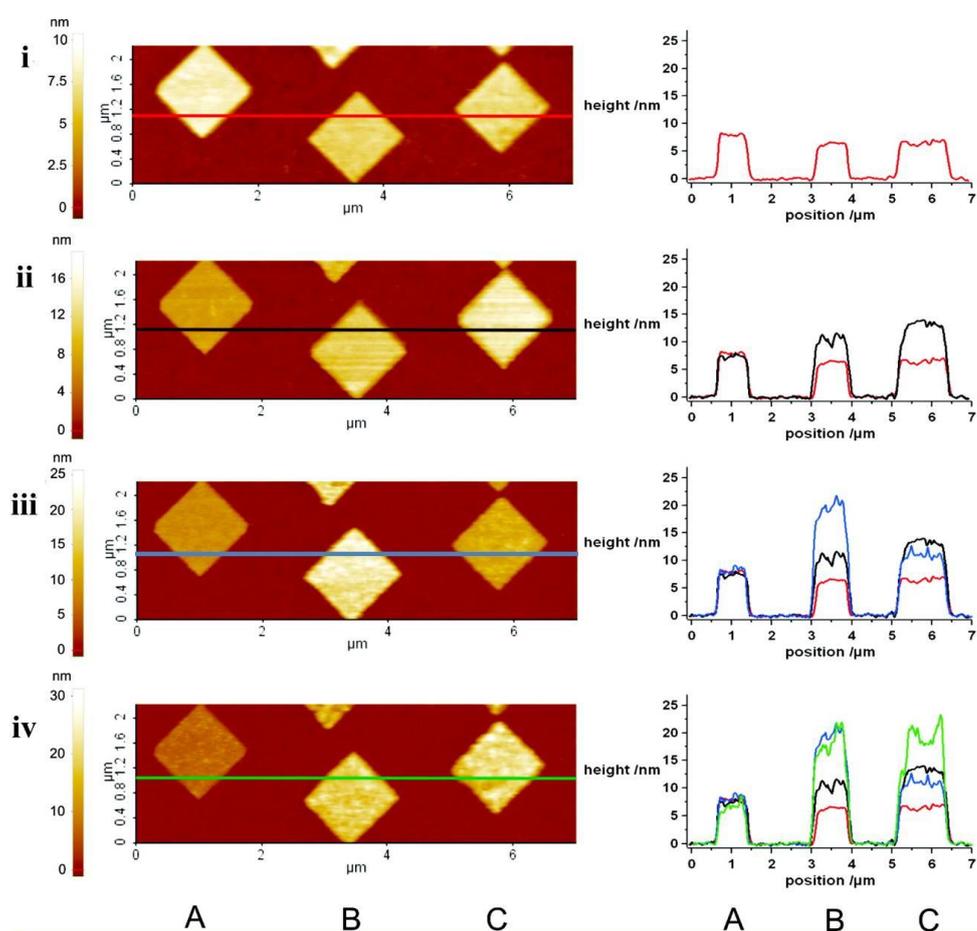
Developing high throughput screening platforms with protein as bio-sensing element is of great importance in experimental biology, biophysics, development of new drugs and the point-of-care diagnostics for profiling proteins that are progressive signature of disease, such as cancer, cardiovascular disease, and so on<sup>82-85</sup>. However, the advances in protein microarrays have been slower due to the challenges associated with protein immobilization. These challenges are; random conformation and orientation after immobilization step<sup>86,87</sup>, difficulty in modifying the protein for proper immobilization on the surface without influencing its bioactivity and native state<sup>85,88</sup>, and short shelf life of the arrays due to protein denaturation<sup>88</sup>. This brings us to another important feature of DNA nanotechnology, which is precise positioning and spatial orientation control of the arrayed biomolecules, via Watson-Crick recognition<sup>11,89,90</sup>. This provides the opportunity to synergistically combine the inherent capacity of DNA and the unlimited functionalities of the protein, to form multi-facet, semi-synthetic DNA conjugates for broad application in bio-theranostics<sup>89,90</sup>.

In the recent review by Niemeyer<sup>18</sup>, the author highlighted the various approaches for crosslinking for synthetic oligonucleotides to protein, via the covalent and non-covalent coupling approaches, these are; biotin–streptavidin interaction<sup>91,92</sup>, Ni-NTA–hexahistidine interaction<sup>93</sup>, reconstitution of apoenzymes<sup>94</sup>, disulphide and maleimide coupling<sup>95,96</sup>, and bio-orthogonal chemistry<sup>97</sup>. These DNA conjugates have been widely used in construction of protein microarrays<sup>89,90</sup> via DNA-directed immobilization (DDI).

The concept of DNA-directed immobilization (DDI) was developed by Christof .M. Niemeyer<sup>98</sup>, which takes the advantage of Watson-Crick hydrogen bonding to immobilize DNA-protein conjugates onto a prior DNA-complementary matrices on the solid support. Some of the advantages of DDI over automated deposition techniques are provision of chemically mild process for parallel immobilization of the protein without instability or denaturation of the protein, reversibility and site selectivity, and spatial orientation without perturbing the bio-activity of the protein.

In recent time, Bano and coworkers<sup>91</sup> synergistically combined nanografting<sup>20,58</sup> and DDI to create spatially defined DNA-protein arrays, with inherent binding capacity to interact with the protein of interest. This approach combined with AFM height

measurements were used to develop protein arrays with nano-molar binding affinity for their specific targets. For example, the authors demonstrated the feasibility and specificity of this synergetic approach, by generating a 3 X 1 array, represented as A, B, and C (see Fig. 1.11 i), where each spot contains distinct ssDNA molecules. Next, the ssDNA matrices were hybridized to complementary DNA-conjugates, which yields protein array; A= control DNA patch without protein, B= DNA-streptavidin, and C= DNA-glucose oxidase. AFM height measurements at the end of this step clearly show an increase in height that detect the successful binding interaction (see Fig. 1.11 ii). Subsequently, the patch B and patch C bind to two different antibodies, anti-STV IgG and anti- GOx IgG, with high selectivity and specificity, with absence of cross reactivity and non-specific interaction (see Fig. 1.11 iii-iv).



**Figure 1.11:** Combination of DDI and nanografting. (i) shows the AFM topographic micrograph of three distinct DNA matrices, A, B, and C, together with the height measurements, corresponding to the red line across the micrograph. The ssDNA arrays were translated to protein array via DDI, as they were incubated with ssDNA-streptavidin conjugates (complementary to patch B), and ssDNA-GOx (complementary to patch C). The AFM height profile in second row (ii, black line) confirms the successful interaction with increase in height. Subsequently, the array was incubated with anti-

STV IgG and anti- GOx IgG in human serum, and its successful recognition is established by the height profile in Fig. 1.11iii (blue line), iv (green line). Copyright © 2009 American Chemical Society.

Also, the methods employed by Bano *et al.* can be extended to other molecular recognition elements such as peptides. In recent times, peptides are emerging as versatile biomolecules whose function can be tailored to a specific molecular function and recognition (e.g. drugs<sup>99,100</sup>, binder<sup>101</sup>, etc.), by combining different 21 amino acids, with high selectivity and specificity<sup>102,103</sup>. Conventionally, peptides exist in nature as hormone peptides, such as oxytocin, vasopressin, motilin, insulin and so on<sup>100,104</sup>. However, peptides can also be designed and developed through genetic and synthetic methods, these are; phage display libraries by random selection<sup>105-107</sup>, intein-based cyclic peptides<sup>100</sup>, miniaturization of known natural receptor down to the synthesizable binding motifs<sup>102,108</sup>, and combination of few amino acids that are known to bind non-covalently to the target of interest<sup>102</sup>. Some of such peptides derived from the aforementioned sources have shown high affinity and selectivity towards their target of interest<sup>105,107,109</sup>. Additionally, peptides are commercially available through solid phase synthesis at a relatively low cost, with high purification and they can be modified with different groups and functionalized with other elements, such as fluorophores, nanoparticles, or affinity tags<sup>110</sup>. This gives them an edge over antibodies, whose production is expensive and tediously difficult for scalability<sup>111,112</sup>.

Out of the many modification strategies<sup>113</sup>, cyclization enhances the mechanical stability by lowering the entropy of the peptides, as such improving the specificity of the peptide towards its target of interest<sup>100,111,114</sup>. This modification is a chemical process that uses the cysteine residues on both N- and C- terminus of the peptide, to transform the peptide from linear configuration to the cyclic configuration<sup>100,113</sup>. Recently, computational-based designs also called *in silico* designs, have aided the design, generation of peptide sequence, and optimization of its tertiary structure towards the binding sites of the target of interest<sup>115,116</sup>. For example, Laio and co-workers<sup>101</sup> employed a combination of molecular dynamics (MD), semi flexible docking, and replica exchange Monte Carlo (MC), to design and develop decapeptide, that is capable of binding small organic molecules like Efavirenz, an antiretroviral medication use for the treatment and prevention of HIV/AIDS. The binding affinity of the designed decapeptide and its interactions towards Efavirenz was experimentally determined using fluorescence spectroscopy and nuclear magnetic resonance spectroscopy (NMR). Their results showed that the binding affinity of the

computationally designed decapeptide towards efavirenz is 64 nM with binding energy of -9.7 kcal/mol.

Inspired by the successful results of Laio *et al.*, Fortuna and her co-workers<sup>111</sup> employed the same algorithm based approach to develop and design eleven distinct octapeptides that show relatively high binding affinity towards the maltose-binding pocket in the maltose binding protein. The authors further validate the binding capacity and the stoichiometry ratio associated with each peptide towards the maltose site on maltose-binding protein (MBP) via electrospray ionisation mass spectrometry (ESI-MS), fluorescence spectroscopy and surface plasmon resonance (SPR). Their results showed that two out of the eleven octa-peptides have their binding affinity towards MBP in the range of micro-molar with one-to-one stoichiometry. Their results were in agreement with the typical binding affinities obtained even with peptides designed and generated through phage display<sup>117,118</sup>.

To push the affinity of peptides towards the detection limit range of antibodies (femtomolar or picomolar), a more strategic approach is to join two low-affinity peptides using a flexible linker, in which each peptide binds to a specific binding site within the protein of interest. The affinity of the coupled binders is predicted to increase nonlinearly, and in turn, targeting two non-equivalent binding sites will lead to higher selectivity. The idea is (1) to generate peptide binders with enhanced binding affinity starting from sequences known to reach  $\mu\text{M}$  affinities towards a target by selecting them either experimentally or computationally; and (2) to choose that best peptides that maximize the attainable binding affinity by carefully characterizing the performance of the shortlisted candidates. This type of approach has been employed to develop high affinity ligands to aid the discovery of drugs<sup>119</sup> and inhibitors<sup>120</sup>. For example, Shuker *et al.*<sup>119</sup> reported an approach based on structural-activity-relationship (SAR) anchored on nuclear magnetic resonance (NMR) spectroscopy, through which the authors could identify several ligands, in which each individual ligand bind to protein subsite pocket with affinity ranging from 2  $\mu\text{M}$  to 1 mM, as measured with fluorescence-based assays. Subsequently, the authors chemically linked five pairs of ligands, starting from weakly bound using from a set of 10 elemental components, with each individual component recognizing a different binding pocket of a protein. Their findings showed that the coupled-binders recognize the immunosuppressant FK506 binding protein (FBP), with high affinities ranging from of 19 nM to 228 nM, with the latter achieved by coupling components that individually reached binding affinities not higher than mM. Despite these results, the idea of generating

nanosensors comprised of coupled-ligands designed with computational algorithms, and later immobilized on surfaces using nanografting and DNA-directed immobilization have not been reported yet in the literature, and it is one objective of the work undertaken in this PhD thesis.

Encouraged and inspired by the versatility and functionalities of *in-silico* based short peptides in biosensing and nanomedicine, the second part of my thesis is aimed at; (1) developing a label-free and smart sensing strategy that combines nanografting, DNA-directed immobilization, and AFM differential height measurements, to immobilize computation-evolved, cyclic peptides on ultra-flat gold substrates with spatial orientation, and able to recognize biomarkers that are overexpressed in ovarian cancer, i.e. Beta-2-microglobulin ( $\beta$ 2m); (2) To produce and characterize two distinct peptide nanoarrays using the synergetic approach in (1), with each probe being comprised of a unique 10-amino acid long cyclic peptide capable of binding to different solvent-exposed binding site on the  $\beta$ 2m with high affinity. The optimum approach for achieving the second objective is to couple the two peptides, one for each binding site, and measure the binding affinity increase in a label-free manner. In an ideal scenario by coupling two binders we would expect their binding free energies to sum up (i.e.  $\Delta G_{TOT} = \Delta G_1 + \Delta G_2$ ), leading the coupled system to reach pM affinities (i.e.  $K_D = \exp [\ln (k_{D1}) + \ln (k_{D2})]$ ), by recalling that  $\Delta G = -RT \ln (k_D)$ .

## 1.6 Thesis outline and contributions

The succeeding chapters in this thesis focus on our application of DNA nanotechnology to tune the function of DNA-binding protein, set-up a programmable peptide arrays for detection of biomarkers and finely-control of the supramolecular assembly of DNA nanostructures via pH.

In the chapter 2, we will describe the function and recognition of site-specific DNA-binding protein within highly-dense spatially-resolved and sequence-dependent DNA nanoreactors. We employ both nanografting and spontaneous formation of self-assemble monolayer (SAM) of pre-hybridized double stranded (ds)DNA, to construct the DNA nanoreactors with varying inherent heterogeneity and homogeneity molecular structure of the DNA nanoreactors. We will also describe the discovery that we obtained, which show the possibility to fine-tune and control the recognition of a DNA endonuclease, namely the homodimeric protein complex termed BamHI, in both cognate and non-cognate nanostructures.

In chapter 3, we report the feasibility of using computer-based algorithms to design, generate and screen highly specific cyclic peptides, of 10-amino acids long, capable of binding to solvent-exposed site on the protein of interest (POI), which is in our case Beta-2-microglobulin ( $\beta$ 2m). Our collaborator on computer-based, Dr Miguel Soler and Dr Sara Fortuna, utilized computation approach based on optimized molecular dynamics and optimized Monte Carlo replica exchange to design peptides with single disulphide bridge for cyclization. Besides, the computer-based screening, another set of screening were performed using SPR characterization technique, in collaboration with researcher at the University of Naples Federico II, Italy. Based on the screening results, we chose pep461, because of it high scoring energy and signal/noise ratio. Then we set up an experimental validation approach, which synergizes nanografting, DNA-directed immobilization, and AFM differential height measurements to immobilize, construct spatially oriented DNA-pep461 arrays, that recognise  $\beta$ 2m and detect the binding interaction through the AFM topographic height measurements. Our results pinpoint the dependency of the  $\beta$ 2m recognition on the density of the DNA-peptide arrays. Also, we discover that pep461 sensitivity is higher ( $\sim 80 \mu\text{M}$ ) than the value obtained with SPR analysis ( $\sim 400 \mu\text{M}$ ).

Subsequently in chapter four, we report an experimental validation of *in silico* peptides, designed by our collaborators, Dr Miguel Soler and Dr Sara Fortuna through a new computer-based algorithm, which allows the identification and generation of possible binding sites on  $\beta$ 2m via webtool Peptimap. This approach is different from the one used in chapter 3. With such computation strategies, pep331 was selected for experimental validation. Afterward, we selected one peptide (pep381) from the pool of peptides designed by our collaborators in chapter 3, and the pep331 generated using the new algorithm (based on webtool Peptimap), and we set up an AFM based experimental validation, to determine the sensitivity of both peptides. Additionally, these two peptides have different amino-acid sequences (each with 10 amino-acids long with single disulphide bridge) and were designed to bind to two distinct solvent exposed binding sites on  $\beta$ 2m. Through the label free AFM experimental approach (same as set up in chapter 3), we discover that the two peptides, pep331 and pep381, have sensitivity of  $\sim 7 \mu\text{M}$  towards  $\beta$ 2m, a higher order sensitivity compared to peptide 461 (tested in chapter 3), which has sensitivity towards  $\beta$ 2m at  $80 \mu\text{M}$  of the target  $\beta$ 2m solution. This report is a proof of concept results, aiming at designing surface-bound bidentate synthetic receptors for biomarker detection.

In chapter 5, we summarize our discoveries and highlight the possible directions with respect to each experiment work in this thesis.

Appendix A reports the last experimental work of this thesis where we demonstrate a strategy to control the supramolecular assembly of DNA nanostructures using pH as trigger. This strategy is based on a novel rational, strand-displacement approach that responds to pH changes and activates a downstream DNA tile self-assembly process. We will describe a switch-like behaviour in the formation of DNA tile-based chains (or tubes), at basic isothermal conditions. Here, the tile-based chains fully form within 24 hours of assembly time, while at acidic conditions no tile chain is formed. The work was carried out in collaboration with Dr. Alessia Amodio and Prof. Francesco Ricci in the Department of Chemistry at University of Rome “Tor Vergata” and co-workers. Alessia performed the fluorescence experiments, while I carried out all the AFM-based designs and experiments.

Appendix B reports the detailed computational-based approach employed by our theoretical group, to design, generate and screen pep381 and pep461, that were validated experimentally in chapter 3 and 4. Also, this section further highlights the additional results obtained by our collaborators (Dr. Anna Russuo, Dr. Daniella Marasco, Ms. Cristina Cantarutti, and Dr. Alessandra Corazza) using SPR, ITC, and NMR to characterize the binding and the interaction between  $\beta$ 2m and cyclic peptides.

Lastly, appendix C describes the details of the new computational approach based on Fast-Fourier Transform molecular docking, employed by Dr Miguel Soler and Dr Sara Fortuna, which is based on primitive identification of binding sites on  $\beta$ 2m, an algorithm driven by Fast-Fourier Transform. This approach was applied to the development of pep331 while the inherent methods for screening and scoring are also reported.

All the AFM-based experiments in this thesis were carried out in molecular nanotechnology for life science applications (MONALISA) laboratory in the Department of Medical and Biological Sciences at University of Udine, Udine, Italy. The remarkable results produced in this PhD thesis have yielded two publications and other two manuscripts in preparation, which will be sent for publication very soon;

- **A.F. Adedeji**, D. Choi, V. Inverso, S.K. Redhu, M. Vidonis, L. Crevatin, A.W. Nicholson, G. Scoles, M. Castronovo. On-off recognitions of endonucleases with DNA brushes of variable density, In preparation.
- M. A. Soler, A. Rodriguez, A. Russo, **A.F. Adedeji**, C. J. D. Founthum, C. Cantarutti, E. Ambrosetti, L. Casalis, A. Corazza, G. Scoles, D. Marasco, A. Laio and S. Fortuna. Computational design of cyclic peptides for the customized oriented immobilization of globular proteins, *Phys. Chem. Chem. Phys.*, 2017, Advance Article, DOI: 10.1039/C6CP07807A.
- **A. F. Adedeji**, M. Soler, D. Marasco, G. Scoles, M. Castronovo, S. Fortuna. Quantitative and label-free characterization of binding affinity of computationally-designed bidentate peptide for detection and immobilization of protein on surfaces using atomic force microscope, In preparation.
- A. Amodio, **A.F. Adedeji**, M. Castronovo, E. Franco, F. Ricci. pH-controlled assembly of DNA tiles. *Journal of the American Chemical Society*, 2016, 138 (39), 12735–12738. **DOI:** 10.1021/jacs.6b07676.

## 1.7 References

1. Nucleic acid structure. at <<http://www.atdbio.com/content/5/Nucleic-acid-structure>>
2. Moeinzadeh, S. & Jabbari, E. *Handbook of Nanomaterials Properties. Handbook of Nanomaterial Properties* (2014). doi:10.1007/978-3-642-31107-9
3. Sun, Y. & Kiang, C. *DNA-based Artificial Nanostructures : Fabrication , Properties , and Applications. 2*, (2005).
4. Watson James. in 1–33
5. Lane, A. M. Original Papers. **44**, 130–136 (2013).
6. Woo, S. Beyond Watson and Crick : Programming the Self - Assembly and Reconfiguration of DNA Nanostructures Based on Stacking Interactions. **2013**, (2013).
7. Reif, J. & Labean, T. Self-assembled DNA Nanostructures and DNA Devices. *NanoFabrication Handb.* 299–328 (2012). doi:doi:10.1201/b11626-17  
10.1201/b11626-17
8. DNA Pearson education.
9. Watson, J. D. & Crick, F. H. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature* **171**, 737–8 (1953).
10. Hood, L. & Galas, D. The digital code of DNA. *Nature* **421**, 444–448 (2003).
11. Pinheiro, A. V, Han, D., Shih, W. M. & Yan, H. Challenges and opportunities for structural DNA nanotechnology. *Nat. Nanotechnol.* **6**, 763–72 (2011).
12. Rothmund, P. W. K. Folding DNA to create nanoscale shapes and patterns. *Nature* **440**, 297–302 (2006).
13. Karzbrun, E., Tayar, A. M., Noireaux, V. & Bar-Ziv, R. H. Synthetic biology. Programmable on-chip DNA compartments as artificial cells. *Science* **345**, 829–32 (2014).
14. Buxboim, A., Daube, S. S. & Bar-Ziv, R. Synthetic gene brushes: a structure–function relationship. *Mol. Syst. Biol.* **4**, 1–8 (2008).

15. Bracha, D., Karzbrun, E., Shemer, G., Pincus, P. a & Bar-Ziv, R. H. Entropy-driven collective interactions in DNA brushes on a biochip. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 4534–8 (2013).
16. Bar, M. & Bar-Ziv, R. H. Spatially resolved DNA brushes on a chip: Gene activation by enzymatic cascade. *Nano Lett.* **9**, 4462–4466 (2009).
17. Bandy, T. J. *et al.* DNA as supramolecular scaffold for functional molecules: progress in DNA nanotechnology. *Chem. Soc. Rev.* **40**, 138–148 (2011).
18. Niemeyer, C. M. Semisynthetic DNA-protein conjugates for biosensing and nanofabrication. *Angew. Chemie - Int. Ed.* **49**, 1200–1216 (2010).
19. Reif, J. H. & LaBean, T. H. DNA Nanotechnology and its Biological Applications. *Bio-Inspired Nanoscale Integr. Comput.* 349–375 (2009).  
doi:10.1002/9780470429983.ch14
20. Mirmomtaz, E. *et al.* Quantitative study of the effect of coverage on the hybridization efficiency of surface-bound DNA nanostructures. *Nano Lett.* **8**, 4134–4139 (2008).
21. Castronovo, Matteo, Radovic slobodanka, grunwald christian, Morgante michella, casalis loredana and S. G. Control of Steric Hindrance on Restriction Enzyme Reactions with Surface-Bound DNA Nanostructures. *Nano Lett.* **8**, 4140–4145 (2008).
22. Castronovo, M. *et al.* Two-dimensional enzyme diffusion in laterally confined DNA monolayers. *Nat. Commun.* **2**, 297 (2011).
23. Steel, a B., Levicky, R. L., Herne, T. M. & Tarlov, M. J. Immobilization of nucleic acids at solid surfaces: effect of oligonucleotide length on layer assembly. *Biophys. J.* **79**, 975–981 (2000).
24. Herne, T. & Tarlov, M. Characterization of DNA probes immobilized on gold surfaces. *J. Am. Chem. Soc.* **7863**, 8916–8920 (1997).
25. Steel, A. B., Herne, T. M. & Tarlov, M. J. Electrochemical Quantitation of DNA Immobilized on Gold. **70**, 4670–4677 (1998).
26. Nkoua Ngavouka, M. D., Bosco, A., Casalis, L. & Parris, P. Determination of

- average internucleotide distance in variable density ssDNA nanobrushes in the presence of different cations species. *Macromolecules* **47**, 8748–8753 (2014).
27. Redhu, S. K., Castronovo, M. & Nicholson, A. W. Digital imprinting of RNA recognition and processing on a self-assembled nucleic acid matrix. *Sci. Rep.* **3**, 2550 (2013).
  28. Bracha, D., Karzbrun, E., Daube, S. S. & Bar-Ziv, R. H. Emergent properties of dense DNA phases toward artificial biosystems on a surface. *Acc. Chem. Res.* **47**, 1912–1921 (2014).
  29. Buxboim, A. *et al.* A single-step photolithographic interface for cell-free gene expression and active biochips. *Small* **3**, 500–510 (2007).
  30. Piner J.; Xu, F.; Hong, S.; Mirkin, C. A., R. D. . Z. Dip Pen Nanolithography. *Science (80-. )*. **283**, 661–663 (1999).
  31. Xu, S. & Liu, G. Y. Nanometer-scale fabrication by simultaneous nanoshaving and molecular self-assembly. *Langmuir* **13**, 127–129 (1997).
  32. Nimse, S. B. alasaheb, Song, K., Sonawane, M. D. igambar, Sayyed, D. R. afiq & Kim, T. Immobilization techniques for microarray: challenges and applications. *Sensors (Basel)*. **14**, 22208–22229 (2014).
  33. Lemeshko, S. V, Powdrill, T., Belosludtsev, Y. Y. & Hogan, M. Oligonucleotides form a duplex with non-helical properties on a positively charged surface. **29**, 3051–3058 (2001).
  34. © 199 6 Nature Publishing Group <http://www.nature.com/naturegenetics>.
  35. Woo, S. & Rothmund, P. W. K. Self-assembly of two-dimensional DNA origami lattices using cation-controlled surface diffusion. *Nat. Commun.* **5**, 4889 (2014).
  36. Abu-Salah, K. M., Ansari, A. A. & Alrokayan, S. A. DNA-based applications in nanobiotechnology. *J. Biomed. Biotechnol.* **2010**, (2010).
  37. Jung, D. *et al.* Covalent Attachment and Hybridization of DNA Oligonucleotides on Patterned Single-Walled Carbon Nanotube Films. 8886–8891 (2004).
  38. Peng, H. *et al.* Label-free electrochemical DNA sensor based on functionalised conducting copolymer. *Biosens. Bioelectron.* **20**, 1821–1828 (2005).

39. Chrisey, L. a *et al.* Covalent attachment of synthetic DNA to self-assembled monolayer films. *Nucleic Acids Res.* **24**, 3031–9 (1996).
40. Pal, S., Kim, M. J. & Song, J. M. Quantitation of surface coverage of oligonucleotides bound to chip surfaces: a fluorescence-based approach using alkaline phosphatase digestion. *Lab Chip* **8**, 1332–1341 (2008).
41. Fixe, F., Dufva, M., Telleman, P. & Christensen, C. B. V. Functionalization of poly(methyl methacrylate) (PMMA) as a substrate for DNA microarrays. *Nucleic Acids Res.* **32**, e9 (2004).
42. Le Berre, V. *et al.* Dendrimeric coating of glass slides for sensitive DNA microarrays analysis. *Nucleic Acids Res.* **31**, e88 (2003).
43. Preininger, C., Bodrossy, L., Sauer, U., Pichler, R. & Weilharter, A. ARChip Epoxy and ARChip UV for covalent on-chip immobilization of pmoA gene-specific oligonucleotides. *Anal. Biochem.* **330**, 29–36 (2004).
44. Gajovic-Eichelmann, N., Ehrentreich-Förster, E. & Bier, F. F. Directed immobilization of nucleic acids at ultramicroelectrodes using a novel electro-deposited polymer. *Biosens. Bioelectron.* **19**, 417–422 (2003).
45. Benters, R., Niemeyer, C. M., Drutschmann, D., Blohm, D. & Wöhrle, D. DNA microarrays with PAMAM dendritic linker systems. *Nucleic Acids Res.* **30**, E10 (2002).
46. Steel, A. B., Herne, T. M. & Tarlov, M. J. Electrochemical quantitation of DNA immobilized on gold. *Anal. Chem.* **70**, 4670–4677 (1998).
47. Chivers, C. E., Koner, A. L., Lowe, E. D. & Howarth, M. How the biotin-streptavidin interaction was made even stronger: investigation via crystallography and a chimaeric tetramer. *Biochem. J.* **435**, 55–63 (2011).
48. Thermo Fisher Scientific. Thermo Scientific Avidin-Biotin Technical Handbook Table of Contents. (2009).
49. Sosnowski, R. G., Tu, E., Butler, W. F., O’Connell, J. P. & Heller, M. J. Rapid determination of single base mismatch mutations in DNA hybrids by direct electric field control. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 1119–1123 (1997).

50. Schreiber, F. Structure and growth of self-assembling monolayers.pdf. *Prog. Surf. Sci.* **65**, (2000).
51. Weiwei, W., Xianglan, W., Deqian, L. & Hongfei, L. Publication Figures principales : a ) b ) c ) m / z. **24**, 685–689 (2006).
52. Iandte, A., Bain, B. C. D. & Whitesides, G. M. with.
53. Mazzarello, R. *et al.* Structure of a CH<sub>3</sub>S monolayer on Au(111) solved by the interplay between molecular dynamics calculations and diffraction measurements. *Phys. Rev. Lett.* **98**, 5–8 (2007).
54. Cossaro, a *et al.* X-ray diffraction and computation yield the structure of alkanethiols on gold(111). *Science (80-. )*. **321**, 943–946 (2008).
55. Boozer, C. *et al.* DNA Directed Protein Immobilization on Mixed DNA/Oligo(ethylene glycol) Self-Assembled Monolayers for Sensitive Biosensors. *Anal. Chem.* **76**, 12297–12302 (2004).
56. Liu, Q. *et al.* DNA computing on surfaces. *Nature* **403**, 175–179 (2000).
57. Arrabito, G. *et al.* Biochips for cell biology by combined dip-pen nanolithography and DNA-directed protein immobilization. *Small* **9**, 4243–4249 (2013).
58. Liu, M. & Liu, G. Y. Hybridization with nanostructures of single-stranded DNA. *Langmuir* **21**, 1972–1978 (2005).
59. Zhang, G.-J., Tanii, T., Funatsu, T. & Ohdomari, I. Patterning of DNA nanostructures on silicon surface by electron beam lithography of self-assembled monolayer. *Chem. Commun.* 786 (2004). doi:10.1039/b315278b
60. Liu, X. & Zhang, H. Scanning Probe Microscopy-Based Nanofabrication for Emerging Applications. *Mater. Sci.* 770–778 (2007).
61. Castronovo, M. *Crowding Effects on Biochemical Reactions With Surface-Bound Dna.*
62. Bracha, D. & Bar-Ziv, R. H. Dendritic and nanowire assemblies of condensed DNA polymer brushes. *J. Am. Chem. Soc.* **136**, 4945–4953 (2014).
63. Daube, S. S., Bracha, D., Buxboim, A. & Bar-Ziv, R. H. Compartmentalization by directional gene expression. *Eur. Cells Mater.* **20**, 49 (2010).

64. Bracha, D., Karzbrun, E., Daube, S. S. & Bar-ziv, R. H. Emergent Properties of Dense DNA Phases toward Artificial Biosystems on a Surface. (2014).
65. Salaita, K., Wang, Y. & Mirkin, C. A. Applications of dip-pen nanolithography. 145–155 doi:10.1038/nnano.2007.39
66. Rozhok, S., Piner, R. & Mirkin, C. A. Dip-pen nanolithography: What controls ink transport? *J. Phys. Chem. B* **107**, 751–757 (2003).
67. Demers, L. M. Direct Patterning of Modified Oligonucleotides on Metals and Insulators by Dip-Pen Nanolithography. *Science* (80-. ). **296**, 1836–1838 (2002).
68. Christman, K. L., Enriquez-Rios, V. D. & Maynard, H. D. Nanopatterning proteins and peptides. *Soft Matter* **2**, 928 (2006).
69. Hyun, J., Ahn, S. J., Lee, W. K., Chilkoti, A. & Zauscher, S. Molecular Recognition-Mediated Fabrication of Protein Nanostructures by Dip-Pen Lithography. *Nano Lett.* **2**, 1203–1207 (2002).
70. Xu, S., Miller, S., Laibinis, P. E. & Liu, G. Y. Fabrication of nanometer scale patterns within self-assembled monolayers by nanografting. *Langmuir* **15**, 7244–7251 (1999).
71. Liu, M., Amro, N. A., Chow, C. S. & Liu, G. Y. Production of Nanostructures of DNA on Surfaces. *Nano Lett.* **2**, 863–867 (2002).
72. Bano, F., Fruk, L., Sanavio, B., Glettenberg, M. & Casalis, L. Towards Multiprotein Nanochips Using Nanografting and DNA Directed Immobilization of Proteins Supporting Information.
73. Sanavio, B. *et al.* Oriented immobilization of prion protein demonstrated via precise interfacial nanostructure measurements. *ACS Nano* **4**, 6607–6616 (2010).
74. Ganau, M. *et al.* A DNA-based nano-immunoassay for the label-free detection of glial fibrillary acidic protein in multicell lysates. *Nanomedicine Nanotechnology, Biol. Med.* **11**, 293–300 (2015).
75. Lee, C. *et al.* Surface Coverage and Structure of Mixed DNA / Alkyl thiol Monolayers on Gold : Characterization by XPS , NEXAFS , and Fluorescence Intensity Measurements. *Society* **78**, 3316–3325 (2006).

76. Peterson, a W., Heaton, R. J. & Georgiadis, R. M. The effect of surface probe density on DNA hybridization. *Nucleic Acids Res.* **29**, 5163–5168 (2001).
77. McGuffee, S. R. & Elcock, A. H. Diffusion, crowding & protein stability in a dynamic molecular model of the bacterial cytoplasm. *PLoS Comput. Biol.* **6**, (2010).
78. Kuznetsova, I., Turoverov, K. & Uversky, V. *What Macromolecular Crowding Can Do to a Protein.* *International Journal of Molecular Sciences* **15**, (2014).
79. Nakano, S. I., Miyoshi, D. & Sugimoto, N. Effects of molecular crowding on the structures, interactions, and functions of nucleic acids. *Chem. Rev.* **114**, 2733–2758 (2014).
80. Elli, R. J. M a c r o m o l e c u l a r c r o w d i n g : o b v i o u s b u t u n d e r a p p r e c i a t e d. **4**, 597–604 (2001).
81. Watkins, H. M., Simon, A. J., Ricci, F. & Plaxco, K. W. Effects of crowding on the stability of a surface-tethered biopolymer: An experimental study of folding in a highly crowded regime. *J. Am. Chem. Soc.* **136**, 8923–8927 (2014).
82. Sutandy, F. X. R., Qian, J., Chen, C.-S. & Zhu, H. Overview of protein microarrays. *Curr. Protoc. protein Sci.* **Chapter 27**, Unit 27.1 (2013).
83. Kodadek, T. Crosstalk Protein microarrays: prospects and problems. (2001).
84. Ramachandran, N., Srivastava, S. & Labaer, J. Applications of protein microarrays for biomarker discovery. *Proteomics. Clin. Appl.* **2**, 1444–59 (2008).
85. Camarero, J. A. Recent developments in the site-specific immobilization of proteins onto solid supports. *Biopolym. - Pept. Sci. Sect.* **90**, 450–458 (2008).
86. Shen, L., Ulrich, N. W., Mello, C. M. & Chen, Z. Determination of conformation and orientation of immobilized peptides and proteins at buried interfaces. *Chem. Phys. Lett.* **619**, 247–255 (2015).
87. Talapatra, A., Rouse, R. & Hardiman, G. Protein microarrays: challenges and promises. *Pharmacogenomics* **3**, 527–536 (2002).
88. No Title. at <[https://en.wikipedia.org/wiki/Protein\\_microarray](https://en.wikipedia.org/wiki/Protein_microarray)>
89. Niemeyer, C. M. The developments of semisynthetic DNA-protein conjugates. *Trends Biotechnol.* **20**, 395–401 (2002).

90. Niemeyer, C. M. Functional devices from DNA and proteins This review provides an overview of the developments in semisynthetic. **2**, 42–52 (2007).
91. Bano, F. *et al.* Toward multiprotein nanoarrays using nanografting and DNA directed immobilization of proteins. *Nano Lett.* **9**, 2614–2618 (2009).
92. Wacker, R. & Niemeyer, C. M. DDI- $\mu$ FIA - A readily configurable microarray-fluorescence immunoassay based on DNA-directed immobilization of proteins. *ChemBioChem* **5**, 453–459 (2004).
93. Goodman, R. P. *et al.* A facile method for reversibly linking a recombinant protein to DNA. *ChemBioChem* **10**, 1551–1557 (2009).
94. Fruk, L. & Niemeyer, C. M. Covalent hemin-DNA adducts for generating a novel class of artificial heme enzymes. *Angew. Chemie - Int. Ed.* **44**, 2603–2606 (2005).
95. Corey, D. & Schultz, P. Generation of a hybrid sequence-specific single-stranded deoxyribonuclease. *Science (80-. )*. **238**, (1987).
96. Pei, D., Corey, D. R. & Schultz, P. G. Site-specific cleavage of duplex DNA by a semisynthetic nuclease via triple-helix formation. *Proc. Natl. Acad. Sci. U. S. A.* **87**, 9858–62 (1990).
97. Lovrinovic, M. *et al.* Synthesis of protein-nucleic acid conjugates by expressed protein ligation. *Chem. Commun.* 822–823 (2003). doi:10.1039/b212294d
98. Niemeyer, C. M., Sano, T., Smith, C. L. & Cantor, C. R. Oligonucleotide-directed self-assembly of proteins : semisynthetic DNA-streptavidin hybrid molecules as connectors for the generation of macroscopic arrays and the construction of supramolecular bioconjugates. **22**, 5530–5539 (1994).
99. Loffet, A. Peptides as drugs: Is there a market? *J. Pept. Sci.* **8**, 1–7 (2002).
100. Joo, S. H. Cyclic peptides as therapeutic agents and biochemical tools. *Biomol. Ther.* **20**, 19–26 (2012).
101. Hong Enriquez, R. P. *et al.* Designing short peptides with high affinity for organic molecules: A combined docking, molecular dynamics, and Monte Carlo approach. *J. Chem. Theory Comput.* **8**, 1121–1128 (2012).
102. Pavan, S. & Berti, F. Short peptides as biosensor transducers. *Anal. Bioanal. Chem.*

- 402**, 3055–3070 (2012).
103. Gladich, I. *et al.* Designing high-affinity peptides for organic molecules by explicit solvent molecular dynamics. *J. Phys. Chem. B* **119**, 12963–12969 (2015).
  104. No Title. at <[https://en.wikipedia.org/wiki/List\\_of\\_human\\_hormones](https://en.wikipedia.org/wiki/List_of_human_hormones)>
  105. Li, Y. *et al.* Directed evolution of human T-cell receptors with picomolar affinities by phage display. *Nat. Biotechnol.* **23**, 349–354 (2005).
  106. Zhang, Z., Zhu, W. & Kodadek, T. Selection and application of peptide-binding peptides. *Nat. Biotechnol.* **18**, 71–74 (2000).
  107. Tadepalli, S. *et al.* Peptide functionalized gold nanorods for the sensitive detection of a cardiac biomarker using plasmonic paper devices. *Sci. Rep.* **5**, 16206 (2015).
  108. Silvia Pavan. Peptides as recognition and sensing elements.
  109. Nakamura, C., Inuyama, Y., Shirai, K., Sugimoto, N. & Miyake, J. Detection of porphyrin using a short peptide immobilized on a surface plasmon resonance sensor chip. *Biosens. Bioelectron.* **16**, 1095–1100 (2001).
  110. Liu, Q., Wang, J. & Boyd, B. J. Peptide-based biosensors. *Talanta* **136**, 114–127 (2015).
  111. Russo, A. *et al.* In silico generation of peptides by replica exchange monte carlo: Docking-based optimization of maltose-binding-protein ligands. *PLoS One* **10**, 1–16 (2015).
  112. Wang, W. & Woodbury, N. W. Selective protein-peptide interactions at surfaces. *Acta Biomater.* **10**, 761–768 (2014).
  113. Peptide Modifications. at <[http://www.genscript.com/peptide\\_modification.html?src=pullmenu](http://www.genscript.com/peptide_modification.html?src=pullmenu)>
  114. Rubinstein, M. & Niv, M. Y. Peptidic modulators of protein-protein interactions: Progress and challenges in computational design. *Biopolymers* **91**, 505–513 (2009).
  115. Klepeis, B. J. L. & Floudas, C. A. In silico protein design: A combinatorial and global optimization Approach. *SIAM News* **37**, 1–4 (2004).
  116. Theory and applications of computational peptide design.

117. Han, Y. & Kodadek, T. Peptides selected to bind the gal80 repressor are potent transcriptional activation domains in yeast\*.
118. Zhu, W., Williams, R. S. & Kodadek, T. A CDC6 protein-binding peptide selected using a bacterial two-hybrid-like system is a cell cycle inhibitor. *J. Biol. Chem.* **275**, 32098–105 (2000).
119. Shuker, S. B., Hajduk, P. J., Meadows, R. P. & Fesik, S. W. Discovering high-affinity ligands for proteins: SAR by NMR. *Science* (80-. ). **274**, (1996).
120. Maly, D. J., Choong, I. C. & Ellman, J. A. Combinatorial target-guided ligand assembly: identification of potent subtype-selective c-Src inhibitors. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 2419–24 (2000).

## **2.0 Results and discussion on endonuclease recognition within highly dense DNA nanoreactors\***

In this chapter, we focus on the application of DNA nanotechnology to study the nucleic acid-protein interactions within highly dense and crowded nanosystems. Our goal is to elucidate the biomolecular reactions in conditions that are more similar to the cell interior in terms of macromolecular crowding and compartmentalization, which cannot be accurately reproduced in solution due to solubilisation limits of crowding agents<sup>1</sup>. To this end, we design and construct synthetic functional DNA nanoreactors (FDNs) with inherent heterogeneous molecular structure, either via top-down, i.e. using the atomic force nanolithography-based approach termed Nanografting, or bottom-up by using the spontaneous formation of self-assembled monolayers of dsDNA. We further investigate the recognition and the specificity of highly specific and structurally-defined type II restriction enzyme, BamHI, within such an opportunely designed self-assembled nucleic-acid based nanostructures, via atomic force microscopy. Our results show an unprecedented digital action of BamHI that is triggered by the inherent nano-confinement, within nanoreactors containing cognate and non-cognate DNA molecules.

### **2.1 Endonuclease enzymes**

In nature, endonucleases exist as restriction enzymes that take part in restriction-modification activities in the prokaryotic organisms; they practically serve to protect prokaryotic unicellular organism (bacteria and archaea) against the invading viral genome, by recognising and cutting the invading viral DNA at a specific sites (also known as restriction or recognition sites)<sup>2,3</sup>. There are about four types of endonucleases; however, type I and II will be discussed here while other types are beyond the scope of this work. Type I endonucleases are multi subunits, hetero-oligomeric proteins complex that have an inherent capacity to participate in both restriction and modification activities at the same time. They bind and cut DNA at random position far away (~ 1000 bps) from their recognition sites<sup>2,4</sup> (see Fig. 2.0). Examples are EcoKI and EcoR124I.

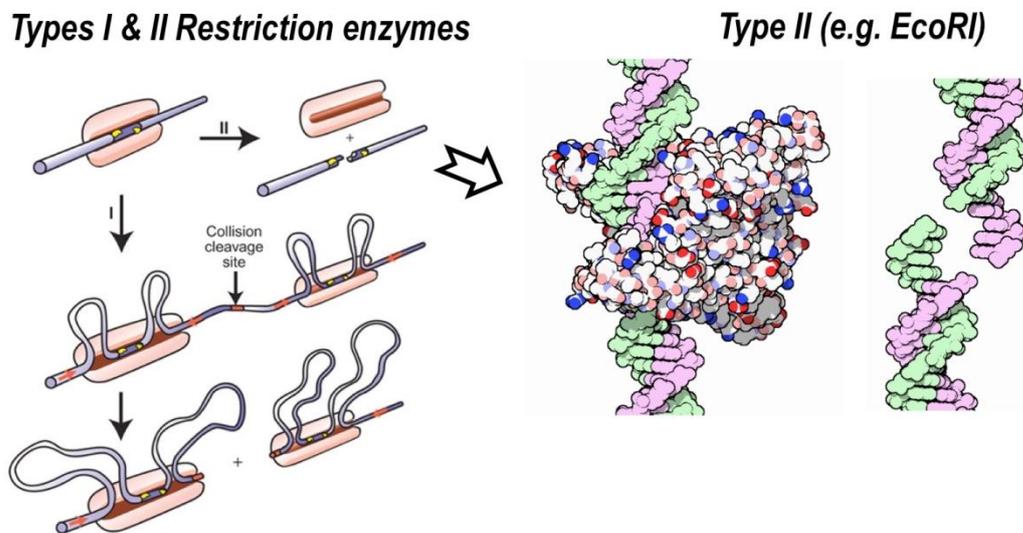
Conversely, type II endonucleases are homo-dimeric protein with separate capacity to serve as either restriction enzymes or modification enzymes but not both. They recognise,

---

\* Manuscript in preparation.

bind and cleave DNA at a well-defined location, either near or within their restriction sites (see Fig. 2.0). Contrast to type I enzymes, they produce discrete DNA fragment visible with distinct gel bands at the completion of their enzymatic activities. Based on their high specificity, they are mostly used in cloning and genetic engineering<sup>5</sup>.

Out of many enzymes in type II category, we work with BamHI, a homodimer protein that recognize short palindromic DNA sequence, which is 6-bps long (5'-GGATCC-3') and cut specifically only at its recognition site<sup>2,6</sup>. In addition to its application in recombinant DNA technologies and genetic engineering, BamHI is widely used as model protein for site-specific protein-DNA interactions<sup>7-9</sup>, because of its stringent and exquisite sequence-specific recognition associated with its affinity for cognate site<sup>2,6,10,11</sup>. However, little is known about the effect of crowding, compartmentalization and confinement on its specificity and selectivity. This is the underlying motivation for this present work.



**Figure 2.0:** Schematic representation of the action of type I and type II endonuclease enzymes. Type II enzymes cleave DNA at a well-defined recognition sites, while Type I cleaves DNA at a random position far from their recognition sites. Adapted from reference 11 and 12.

## 2.2 Design and construction of surface-bound DNA nanoreactors

We designed three distinct, 44-bp-long DNA molecules with recognition sites located in the middle. One is provided with the correct BamHI recognition site (RS) that ensures specificity for the enzyme (termed DNA-1), while the other two is provided with sequences that only partially match such cognate RS (5'-CGATCA-3' and 5'-AGATCA-3', termed DNA-2 and DNA-3 respectively) (see Table 2.0 for the complete sequence). With

such design, we constructed three distinct categories of spatially resolved DNA nanoreactors, with each category made up of unique recognition site on the DNA molecules, either DNA-1 or DNA-2 or DNA-3. These spatially resolved DNA nanoreactors were generated either via top-down, i.e. using Nanografting (see Fig. 2.1), or bottom-up by using the spontaneous formation of self-assembled monolayers (see Fig. 2.2).

Nanografting is a tip-induced self-assembling technique that uses AFM tip to transfer energy onto a selected area at high force ( $\sim 120$  nN), which results into displacement of the existing SAM molecules within the selected area and the immobilization of a new molecules present in the contacting solution<sup>14,15</sup> (which is thiol-modified ssDNA molecules in our case, see Fig. 2.1 a-b). Based on the nanografting parameters developed by Liu<sup>16</sup> and optimized by our group<sup>17</sup>, we finely-controlled the density of ssDNA molecules within the confined section, by varying the number of times the AFM tip over-write the selected area during the nanografting process. At the end of the nanografting process, we carried out hybridization process, which involves the introduction of the complementary ssDNA solution onto the substrate, as shown in (c). Next, we obtained the AFM micrograph of the dsDNA nanoreactor, by imaging the nanografted section at low force ( $\sim 0.2$  nN). Notably, the measured topographic height of such dsDNA nanoreactors is directly related to; 1) the inherent DNA density or, at any fixed density, and 2) to the length of the DNA molecules, as the dsDNA molecules saturates and stand uprightly in the crowded nanostructures<sup>14,15,17</sup>. At the end of the successful generation of DNA nanoreactors, we incubated the various sequence-dependent DNA nanoreactors with  $0.2$  U/ $\mu$ L ( $\sim 20$  nM) BamHI enzyme, at 37-degree Celsius and the standard catalytic reaction lasted for 1 hour. Subsequently, we obtained the AFM topographic height images (at the same low force,  $\sim 0.2$  nN), so as to compare the height profile of each nanoreactor before and after the enzymatic reactions. The measured change in height becomes the signature of the successful recognition of the recognition sites by BamHI. The second approach employed for the construction of crowded DNA nanostructures on an ultra-flat gold chips, will be discussed in part in the next section and in the methods section.

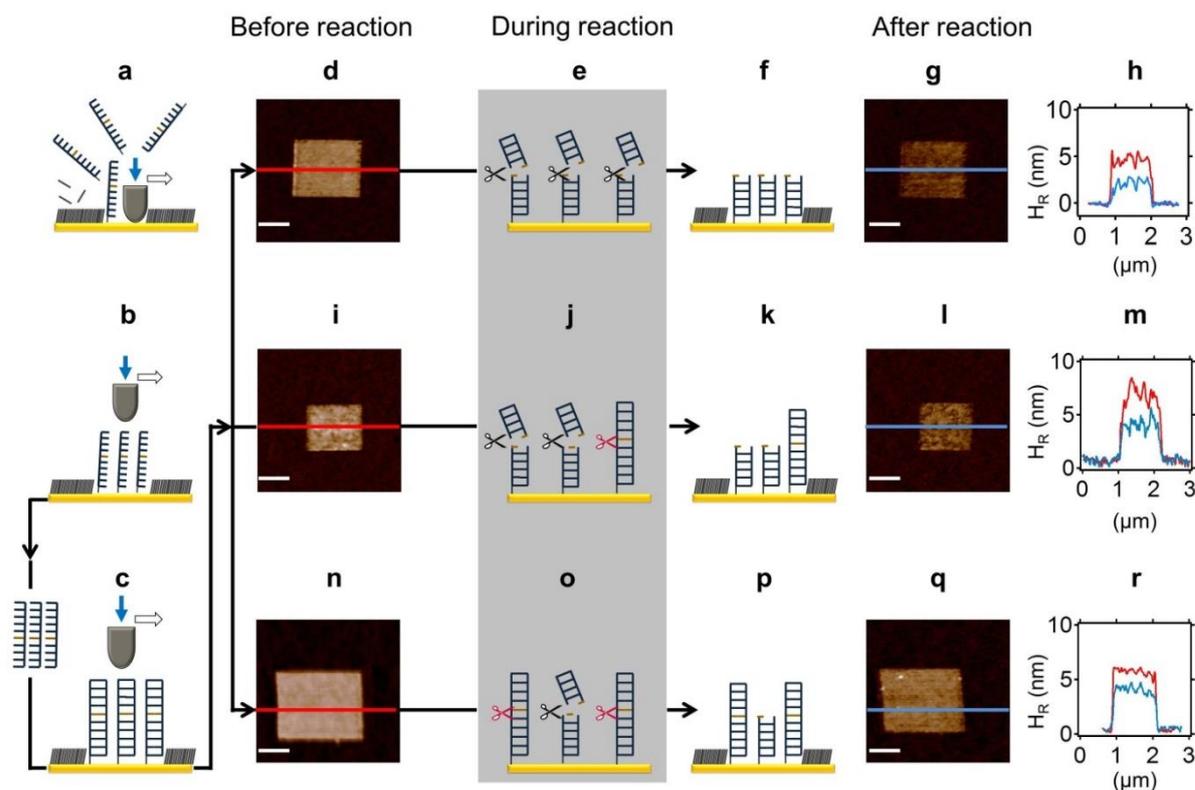
### **2.3 Specificity of BamHI within sequence-dependent DNA nanoreactors**

In the Fig. 2.1, we assessed the specificity of BamHI within sequence-dependent DNA nanoreactors, with fixed dimension  $1 \times 1 \mu\text{m}^2$ , and at the same conditions in term of enzyme concentration, ionic strength, buffering salts and temperature.

The first row of Fig. 2.1 shows the observed behaviour of BamHI within dsDNA-1 molecules. Qualitatively, the colour map of the nanoreactor in the AFM micrograph changed after the enzymatic reaction; brighter signifies increase in height and darker signifies decrease in height. However, to quantify the action of BamHI, we measured the height profile across the confined dsDNA molecules before and after the enzymatic reaction, as shown in (d) and (g). The measured AFM height profiles are shown in (h), which shows that the relative height of the dsDNA-1 (red line) is 5.2 nm and it reduced to 2.5 nm (light blue line) after the enzymatic reaction. Therefore, we can infer that the BamHI successfully recognise the restriction site, located in the middle of the DNA-1 molecules independently of the surface density of DNA-1 molecules within the confined nanostructure. Using the empirical method employed in the previous work of our group<sup>14</sup>, the estimated density corresponding to the relative height of 5.2 nm is 1650 molecules/ $\mu\text{m}^2$ .

Using the same approach of AFM height measurement, we observed an unexpected behaviour in the DNA-2 nanoreactors, represented in the second row of Fig. 2.1. The measured relative height of the dsDNA-2 before enzymatic reaction is 6.5 nm and it reduced to 4 nm after the enzymatic reaction (see (m)). This is an indication BamHI mistakenly cleaved the wrong site, with nearly 50 % reduction in height of the nanoreactors after the enzymatic reaction. The estimated density of DNA-2 molecules is 1900 molecules / $\mu\text{m}^2$ .

For the third mutated sequence, DNA-3, shown in the third row of Fig. 2.1. We observed another anomalous behaviour of BamHI, as the height profile in (r) shows a reduction in height of the nanoreactor, which is low compared to that obtained in DNA-1 (first row) and DNA-2 (second row). The measured relative height of the nanoreactor before enzymatic reaction is 5.5 nm and it becomes 4.5 nm after the enzymatic reactions. This shows that ~25% of the DNA-3 molecules were cleaved by BamHI during the enzymatic reaction, as schematically represented in (2.1p). The estimated density of DNA-3 molecules is 1650 molecules / $\mu\text{m}^2$ , the same as DNA-1.



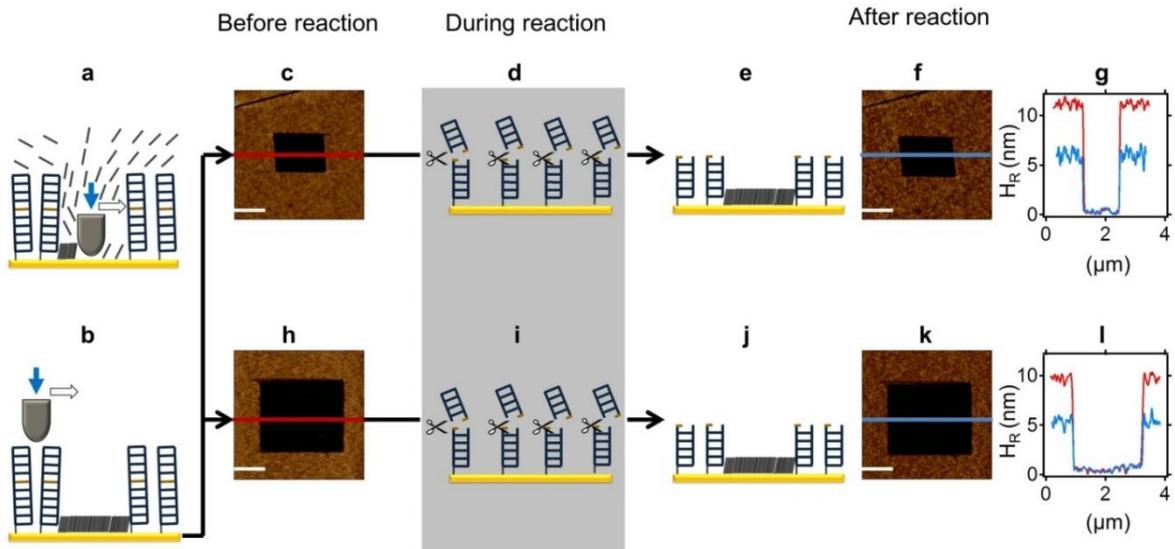
**Figure 2.1:** Generation of three distinct sequence-dependent, spatially resolve DNA nanoreactors, via nanografting. The main procedural steps involve in DNA nanografting are shown in a form of schematics in a-c. In (a), at high applied load ( $\sim 120$  nN) the AFM tip displaces alkyl thiol molecules from the pre-selected area and immobilize ssDNA molecules instead, from the contacting ssDNA solution. At the end of the substitution reaction, the obtained laterally confined ssDNA nanosystem can be imaged using the AFM tip at a low force ( $\sim 0.2$  nN). Next to (b), DNA hybridization is carried out by introducing the solution of complementary ssDNA onto the substrate. Afterwards, the AFM tip ( $\sim 0.2$  nN) is used to image the laterally confined dsDNA nanostructure. The dsDNA molecules confined within the bio-repellent alkyl thiol monolayer consist of restriction site (orange colour line), located at the middle along the length of the molecule. In the first row, d-h represents the action of BamHI observed in nanoreactor made up of DNA-1 molecules (RS-5'-GGATCC-3'). (d) Shows the AFM micrograph of nanografted dsDNA nanoreactor before enzymatic reaction. The schematic in (e) represents the interaction between BamHI (molecular scissors) and dsDNA molecules, resulting into digestion of dsDNA molecules. (f) Shows the schematic representation of the output of enzymatic reaction, in which all the DNA-1 molecules are cleaved in the middle, where the restriction site is located. The same AFM micrograph in (d) is shown in (g) after enzymatic reaction. (h) Shows the AFM height profile that corresponds to the line profile across d (red line) and h (light blue line), indicating the 50% reduction in height of the DNA-1 molecules after enzymatic reaction. The second row, which is i-m, shows the action of BamHI in nanoreactor made up of DNA-2 molecules (RS-5'-CGATCA-3'). (i) Shows the AFM micrograph of laterally confined dsDNA

nanoreactors made up of DNA-2 molecules. (j) Represents the schematic of the interaction between BamHI (molecular scissors) and DNA-2 molecules during enzymatic reaction. The scissors has black colour when the molecules are cleaved and red colour when the molecules are not cut. (k) Depicts the output of enzymatic reaction, in which not all the molecules are cleaved by BamHI. (l) Shows the same AFM micrograph as (i) after enzymatic reaction. (m) Represents the AFM height profile, that corresponds to the line profile (red) across (i) and line profile (light blue) across (l). Lastly, the third row, n-r shows the action of BamHI in nanoreactor made up of DNA-3 molecules (RS-5'-AGATCA-3'). (n) Shows the AFM micrograph of laterally confined dsDNA nanoreactors made up of DNA-3 molecules. (o) Represents the schematic of the interaction between BamHI (molecular scissors) and DNA-3 molecules during enzymatic reaction. (p) Depicts the output of enzymatic reaction, in which not all the molecules are cleaved by BamHI. (q) Shows the same AFM micrograph as (i) after enzymatic reaction. (r) Represents the AFM height profile that corresponds to the line profile (red) across (n) and line profile (light blue) across (q). The scale bar for all the AFM micrograph is 500 nm.

To further characterize and generalize the action of BamHI in crowded DNA nanostructures, we used the second approach to generate crowded DNA nanostructure, which is based on the spontaneous formation of self-assemble monolayer (SAM) of pre-hybridized thiol-derivatised dsDNA molecules, to form micro-scale surface-bound DNA nanostructures (see materials and methods section for details). This approach helps to generate fully-hybridized crowded DNA nanostructures with the highest density range (from 3500 molecules/ $\mu\text{m}^2$  – 5000 molecules /  $\mu\text{m}^2$ ) without compromising the duplex stability<sup>18</sup>. Conversely, as previously established by previous works<sup>17,19</sup> of our group, at the very high density, the hybridization efficiency within nanografted ssDNA patches is only approximately 50%. Therefore, using SAMs circumvents such a setback that is associated with nanografted system at very high density, as such we employed it to reach the high-density range.

With this strategy, we examined the activity of BamHI within sequence-dependent crowded nanostructure with molecular density of 3500 molecules/ $\mu\text{m}^2$ , which correspond to a relative initial height of 10 nm. In the cognate dsDNA SAM, which is the dsDNA monolayer made up of DNA-1, the BamHI cleaved the DNA-1 molecules. This is established by the AFM height measurement of dsDNA SAM before (2.2 c) and after enzymatic reaction (2.2 f). The height profile in (2.2g) shows that the relative initial height (red line) is decreased to half after the enzymatic reactions. This result is in full agreement

with the observed behaviour of BamHI with confined dsDNA nanoreactors made of DNA-1 in Fig. 2.1. Even though the density of DNA-1 (1650 molecules/ $\mu\text{m}^2$ ) in Fig. 2.1 is much lower than the density of DNA-1 in Fig 2.2.



**Figure 2.2:** Construction of highly dense DNA nanoreactors via spontaneous formation of self-assembled monolayer of pre-hybridized dsDNA molecules on ultra-flat gold substrate combined with negative nanografting of ethylene glycol-terminated alkyl thiol, which serves as entrance-passage for BamHI to enter the surrounding dsDNA monolayer. (a) Shows the schematic representation of nanografting of bio-repellent alkyl thiol molecules within the self-assembled monolayer of dsDNA. (b) Depicts the imaging process after (a). The AFM tip scans the self-assembled dsDNA monolayer at a low force ( $\sim 0.2$  nN) with respect the tip-induced self-assembled monolayer of the alkyl thiol. In (c), we have the AFM topographic micrograph of self-assembled monolayer of dsDNA made up of DNA-1, with nanografted monolayer of alkythiol (black-square). (d) Schematic representation of the interaction between the BamHI and the DNA-1 molecules. (e) Depicts the output result after Bam HI activity with DNA-1 molecules, resulting in 50 % reduction in height of DNA-1 molecules. (f) Shows AFM micrograph of the same in (c) after the enzymatic reaction. (g) AFM Height profiles that correspond to the line profiles (red line and light blue line) across (c) and (f) respectively. This indicates that the BamHI activity proceeded to completion as the final height is half of the initial height. The second row show the behaviour of BamHI in self-assembled monolayer made up of DNA-3 molecules. In (h), we have the AFM micrograph of dsDNA monolayer with centralized square shaped ( $2 \times 2 \mu\text{m}^2$  in area) ethylene glycol-terminated alkyl thiol, before enzymatic reactions. Next, (i) depicts the schematic representation of the action of BamHI (black scissors) within the dsDNA monolayer; followed by (j) that represents the output result after the enzymatic reaction. In (k), we have the same AFM micrograph as in (h) after the enzymatic reactions. The height profile in (l) corresponds to the relative height of dsDNA monolayer before (red line) and after (light blue) the enzymatic reactions. The 50 % reduction in the relative

height of DNA-3 molecules after the enzymatic reaction indicates the full digestion of DNA molecules by BamHI, even though the molecules consists of non-cognate restriction sites. The scale bar in all the AFM micrograph is 750 nm.

Conversely, we observed another unexpected behaviour of BamHI in DNA nanostructures made up of DNA-3 (second row of Fig. 2.2), we observed a nearly 50% reduction in the relative height of the DNA-3 after the enzymatic reaction, as shown in the Fig 2.2l. This is different from the action of BamHI within confined dsDNA nanoreactors made up of DNA-3 in Fig. 2.1.

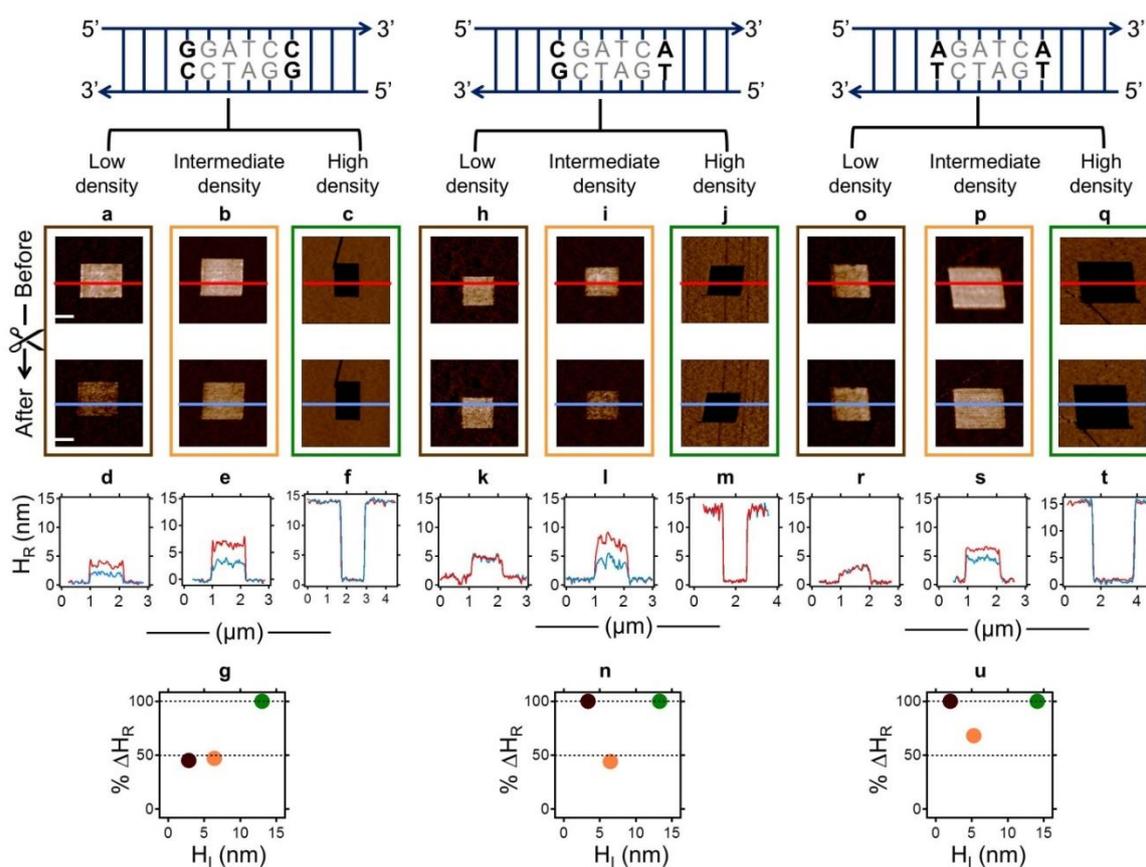
The main question at this point is what is responsible for the anomalous behaviour of BamHI within non-cognate nanoreactors, is it density or confinement or the both, and is this behaviour specific only for a given range of density or it is a uniform observation across the density range, from sparse configuration to the densely-packed configuration. We plan to answer these questions in the next section.

#### **2.4 Sequence-density related action of BamHI within DNA nanoreactors**

In this section, we elucidate the activity of BamHI in sequence-dependent dsDNA nanoreactors, with density ranging from sparse configuration to highly dense configuration, i.e. from 1000 molecules/ $\mu\text{m}^2$  (or less) to 4500 molecules/ $\mu\text{m}^2$ . Using the two approaches described in the previous section: nanografting and spontaneous formation of self-assembled monolayer of dsDNA, we generated three-density base system, where the first and second density regime; low and intermediate density, were generated using nanografting and the high-density region was generated using the SAM approach. In addition, we coupled this density arrangement to the three different DNA molecules (each with distinct recognition site), varying therefore two independent parameters. These partitions (see Fig. 2.3) helps us to systematically follow the mechanistic activity and recognition of BamHI within sequence dependent and density dependent confined DNA nanoreactors.

Starting with cognate DNA nanoreactors, which are represented on the first column of Fig. 2.3, at the low density (a), BamHI recognized the restriction site on the DNA-1 molecules with high specificity, leading to complete digestion of the DNA molecule, as shown by the AFM height profiles in (2.3d), which correspond to the line profiles (red and light blue lines) crossing the AFM micrographs in Fig. 2.3 a. Based on the measured relative heights, the relative height of DNA-1 molecules before BamHI action is 2.8 and it was reduced to 1.3 nm after the enzymatic reaction. Proceeding to the intermediate density, with

1600 molecules/ $\mu\text{m}^2$ , we also observed complete digestion of the DNA-1 molecules by BamHI endonuclease. This is revealed by the AFM height profile in (2.3e), which is the comparison of the relative height ( $H_R$ ) of the AFM micrographs in (2.3b) with respect to before and after enzymatic reactions. However, at the high density (4500 molecules/ $\mu\text{m}^2$ ), with relative height ( $H_R$ ) of 14 nm, BamHI could not digest the highly dense DNA-1 molecules. The possible interpretation for this behaviour of BamHI in density-dependent DNA-1 molecules is that at both low DNA intermediate densities, BamHI still retains its stringent recognition without any perturbation by density. Conversely, at high density, BamHI accessibility into DNA-1 monolayer is hindered by the high level of crowding and steric hindrances within the monolayer, with short inter-dsDNA distances. These results is in full agreement with the published results from our group<sup>14,20</sup>.



**Figure 2.3:** Sequence-density related function of BamHI revealed by AFM height measurements. The first, second and third column, show the observed action of BamHI within the spatially confined dsDNA nanoreactors that is made up of DNA-1, DNA-2 and DNA-3 respectively. The upper section of each column shows the schematic representation of dsDNA molecules with restriction site located in the middle. The core sequence of the recognition site (RS) is written in grey while the edges are written in bold black colour. The main difference in the three restriction sites is located at the edges

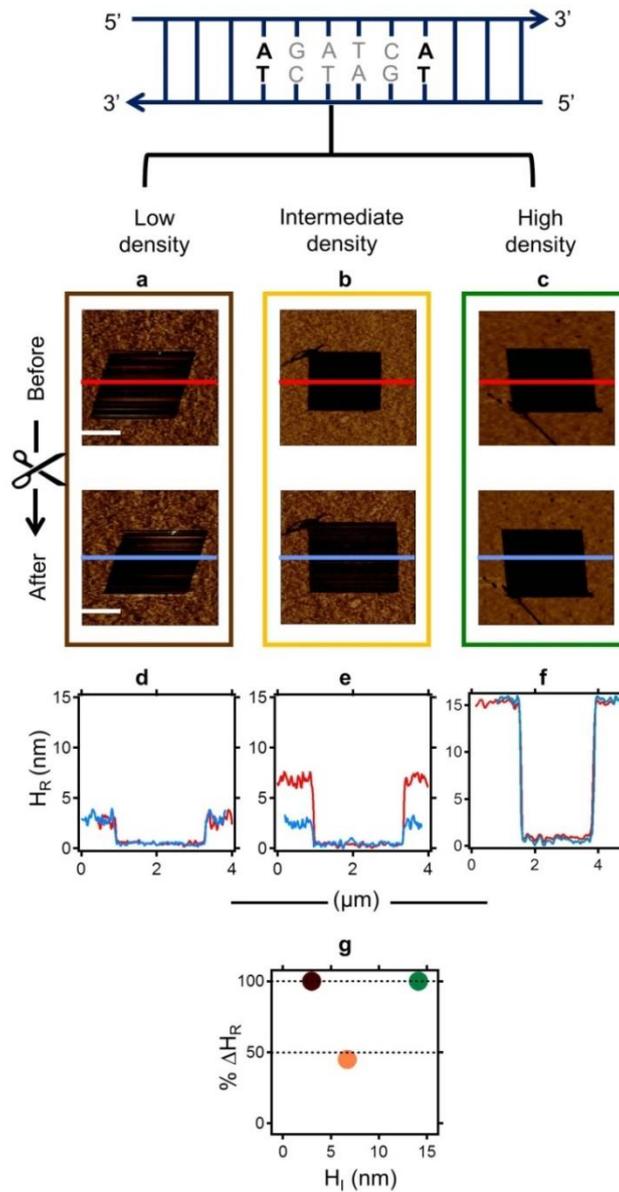
of the recognition sequences while the core (in grey) remains the same for all the three DNA molecules. The density range considered for DNA-1, DNA-2 and DNA-3, is partitioned into three distinct density domains; low density, intermediate density and high density, respectively. In frame (a-c), we have the AFM micrographs of the confined DNA-1 molecules before (up) and after (down) BamHI action. The colour code, brown, orange, and green is designated for low density, intermediate density and high density respectively. Followed by (d-e), which represent the AFM height profiles, measured relatively to the surrounding bio-repellent alkyl thiol monolayer, which correspond to the line profiles (red and light blue) across the AFM micrographs in the low-density frame, intermediate density frame, and high density frame, respectively. (g) Is a plot that shows the relationship between the percentage change in the relative height ( $\Delta H_R$ ) of nanostructures in (a-c), as calculated from (d-e) against the initial height ( $H_I$ ) of the nanoreactors before the enzymatic reaction. We observed 50% reduction in the relative height of confined nanostructures in (a) and (b), while we observed no reduction in height in the nanoreactor in (c). In the second column, (h-j), we have the AFM micrographs of the confined DNA-2 molecules before (up) and after (down) Bam HI action, with respective designated densities. Followed by (k-m), which represent the AFM height profiles, measured relatively to the surrounding bio-repellent alkyl thiol monolayer that corresponds to the line profiles (red and light blue) across the AFM micrographs with respect to the density range. (n) Is the plot that shows the relationship between the percentage change in the relative height ( $\Delta H_R$ ) of nanostructures in (h-j), as calculated from (k-m) against the initial height ( $H_I$ ) of the nanoreactors before the enzymatic reaction. We observed no reduction in the low-density regime, 50% reduction in relative height in the intermediate density regime and no reduction in the relative height at the high-density regime, after the enzymatic reaction. This shows a switch-like activity of BamHI, activated by density and confinement. In the third column, (o-q) which is designated to the low density, intermediate density and high density respectively, we have the AFM micrographs of the confined DNA-3 molecules before (up) and after (down) BamHI action. Next is (r-t), which represents the AFM height profiles, measured relatively to the surrounding bio-repellent alkyl thiol monolayer that correspond to the lines (red and light blue) across the AFM micrographs with respect to the respective density range. (u) Is the digital-box that shows the relationship between the percentage change in the relative height ( $\Delta H_R$ ) of nanostructures in (o-q), as calculated from (r-t) against the initial relative height ( $H_I$ ) of the nanoreactors before the enzymatic reaction. At low density regime, there was no reduction of relative height after the enzymatic reaction; at the intermediate density regime, we observed uncompleted activity of BamHI, with corresponding ~30% reduction in relative height: lastly at the high-density regime, there was no reduction in height, evidently the initial relative height is equal to the final relative height of the DNA-3 molecules after the enzymatic reaction. The scale bar is 500 nm and it applies to all AFM micrograph in this Figure.

Conversely, we observed unprecedented, density-dependent behaviour of BamHI for DNA-2 and DNA-3 molecules, as represented in the second and the third column of Fig. 2.3. For each DNA molecule, we observed three distinct regimes. At low density ( $\sim 1300$  molecules/ $\mu\text{m}^2$ ), BamHI cleaves none of the two. This is shown by the height profiles in (2.3 k, r), which correspond to the lines across the AFM micrographs in (2.3 h, o). The same result is observed in the DNA monolayer of DNA-3 (see Fig. 2.4 d). This is in agreement with its known high specificity for the canonical site (DNA-1) in dilute systems<sup>6,11</sup>. However, at the intermediate density ( $\sim 1700$  molecules/ $\mu\text{m}^2$ ), we observed two distinct behaviours for DNA-2 and DNA-3 molecules. For DNA-2, BamHI digested the partially consensus molecules, as shown by the height profiles in (l). The relative height of the DNA molecules before the enzymatic reaction (red line) is 6.5 nm and it decreased to 2.9 nm after the enzymatic reaction (light blue line). For DNA-3, BamHI digested only a fraction of the molecules within the nanoreactor. This is shown by the AFM height measurement represented in (2.3s), which corresponds to the height profiles (red and light blue) of the AFM micrograph in (2.3p, up and down section). The relative height measured with respect to the surrounding monolayer of bio-repellent ethylene glycol terminated alky thiol is 5.2 nm and it decreased to a relative height of 3.5 nm after the enzymatic reaction (see 2.3s). As shown in (2.3u), the percentage decrease in height after the enzymatic reaction is  $\sim 30\%$ , whereas, the percentage decrease in the relative height for the same molecules (DNA-3) generated by self-assembled monolayer approach, is 50%, as shown in Fig. 2.4g.

Lastly, at high density regime ( $4500$  molecules/ $\mu\text{m}^2$ ), we observed no reduction in relative height, shown in (2.3m, t). The same behaviour is also observed in (2.4f). This shows that BamHI could not gain access into the crowded monolayer because of the volume occupancy in the system.

These results show that the unexpected digestion of Bam HI with varying efficiencies in both DNA-2 and DNA-3 nanoreactors only occurs at the intermediate densities, and in addition is incomplete (with respect to the action in DNA-1 nanoreactors). The incomplete action of BamHI at the intermediate density for DNA-2 of non-cognate DNA nanoreactors, could be explained by the fact that the nanoreactors consists in a heterogeneous ensemble of micro-domains with different densities, as suggested by the surface roughness of the nanopatches. In this sense, the presence of multiple density domains within each of the nanoreactor would make BamHI to be initially triggered at intermediate densities, and finally inhibited as soon as the average density of the unreacted DNA molecules reaches a sufficiently low level. This hypothesis is explored in the next

section by developing a quantitative approach to quantify the heterogeneity of arrangement for all the different dsDNAs in the nanoreactors and its impact on the action of BamHI.



**Figure 2.4:** Sequence-density related activity of BamHI within self-assembled monolayer of dsDNA made up of DNA-3. The upper section shows the schematic representation of dsDNA-3 molecules with restriction site located in the middle. The core sequence of the recognition site (RS) is written in grey while the edges are written in bold black colour. The density range is partitioned into three distinct density domains: low density, intermediate density and high density respectively. In frame (a-c), with designated densities, we have the AFM micrographs of the confined DNA-3 molecules before (up) and after (down) Bam HI action. The colour code, brown, orange, and green is designated for low density, intermediate density and high density respectively. Followed by (d-e), which represent the AFM height profiles, measured relatively to the surrounding bio-repellent alkyl thiol monolayer, which correspond to the line profiles (red and light blue) across the AFM micrographs

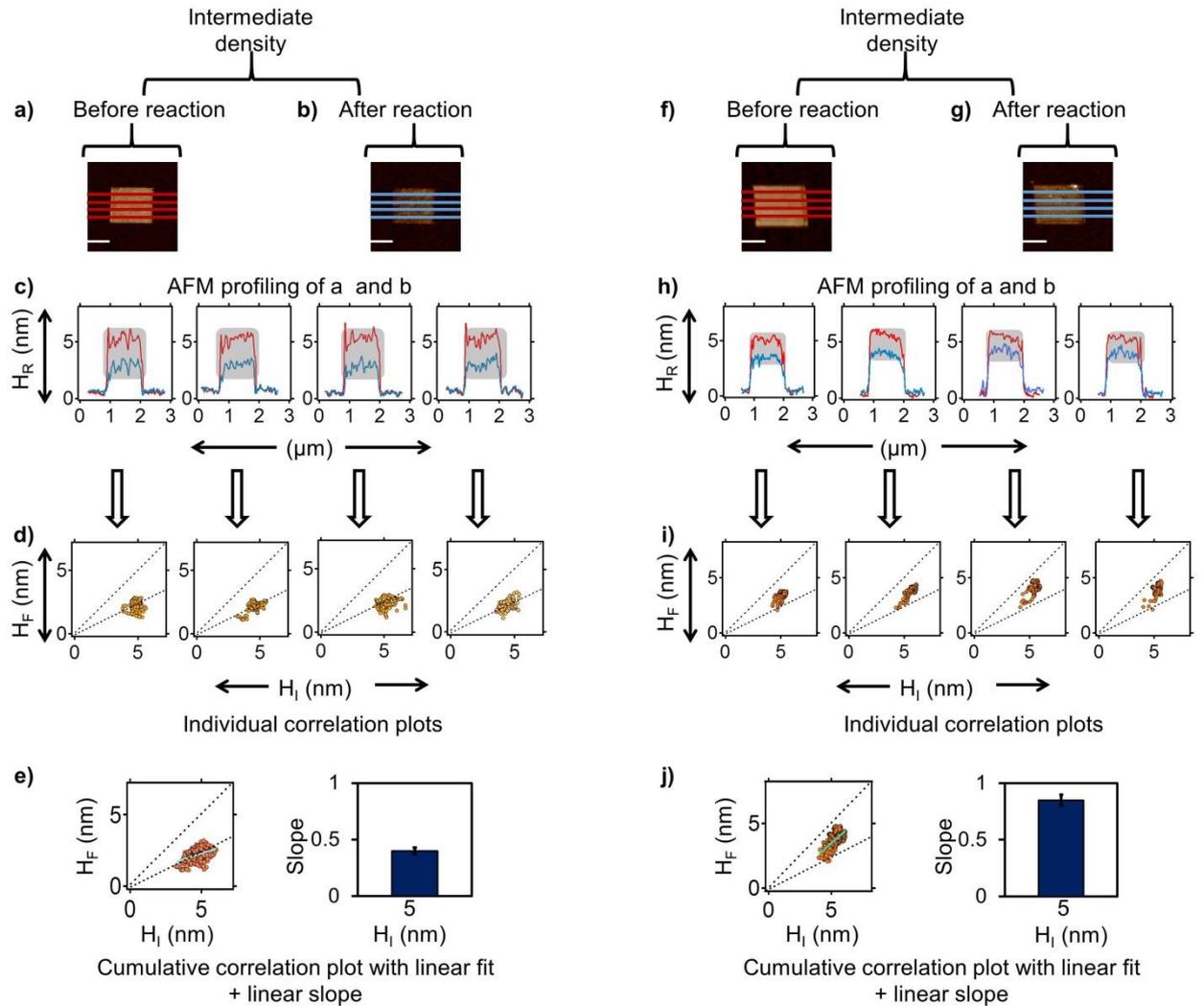
in the low-density frame, intermediate density frame and the high-density frame. (g) Is the plot that shows the relationship between the percentage change in the relative height ( $\Delta H_R$ ) of nanostructures in (a-c), as calculated from (d-e) against the initial height ( $H_i$ ) of the nanoreactors before the enzymatic reaction. The scale bar is 1  $\mu\text{m}$  and it is the same for all the AFM micrograph in this Figure. The enzymatic reaction leads to no height reduction in the low-density regime, 50% reduction in relative height in the intermediate density regime, and no reduction in the relative height at the high-density regime. This results for DNA-3 SAMs shows a switch-like activity of BamHI, activated by density and confinement, which is similar to our findings for DNA-2 nanografted patches.

## **2.5 Tracking and capturing the endonuclease recognition within DNA nanoreactors**

To reveal the effect of nanostructure arrangement of the dsDNA molecules on the activity of BamHI, we developed a side-by-side AFM height profiling, which involves the partitioning of the AFM topographic images that correspond to before and after restriction enzymatic reaction into subunits of nanoreactors (separated by 200 nm in the slow scan direction, Y), followed by the line profile across each subunit (see Fig. 2.5 & 2.6), and then we translated the height measurements into a correlation analysis of the individual height profiles. This analytical approach is intended to help detect different enzyme behaviours in all the different domains within each nanoreactor, avoiding the artefacts caused by height averaging. All the analysis was carried out in IGOR PRO 6.3 workspace.

In Fig. 2.5, we show two different DNA nanoreactors belonging to the intermediate density regime, analysed by the novel side-by-side AFM height measurement approach. The first column and the second column in Fig. 2.5 represent the DNA-1 and DNA-3 nanoreactors with respect to before and after BamHI action. To track the BamHI recognition within the two sequence-dependent nanoreactors, we measured four AFM height profile, each profile across the subunit ( $\sim 200$  nm in y-axis), with one-to-one profile mapping across AFM micrographs, which correspond to before and after the enzymatic reaction (see Fig. 2.5a-b, and f-g). This resulted into eight AFM height profiles, four for DNA-1 (left, c) and four for DNA-3 (right, h), with each AFM profile tallying with the line profiles in Fig 2.5a-b, and f-g. We observed an even action of BamHI within DNA-1 nanoreactors (cognate nanostructure), as revealed by the individual AFM height profiles, with relative final height (light blue) equal to the half of the initial relative height (red line), as shown in Fig. 2.5c. However, the opposite is the case in the DNA-3 nanoreactor (Fig. 2.5h), each individual

AFM height profile shows a varying and unequal action of BamHI with respect to each subunit or sub-density domain within the nanoreactor (see Fig. 2.5i).

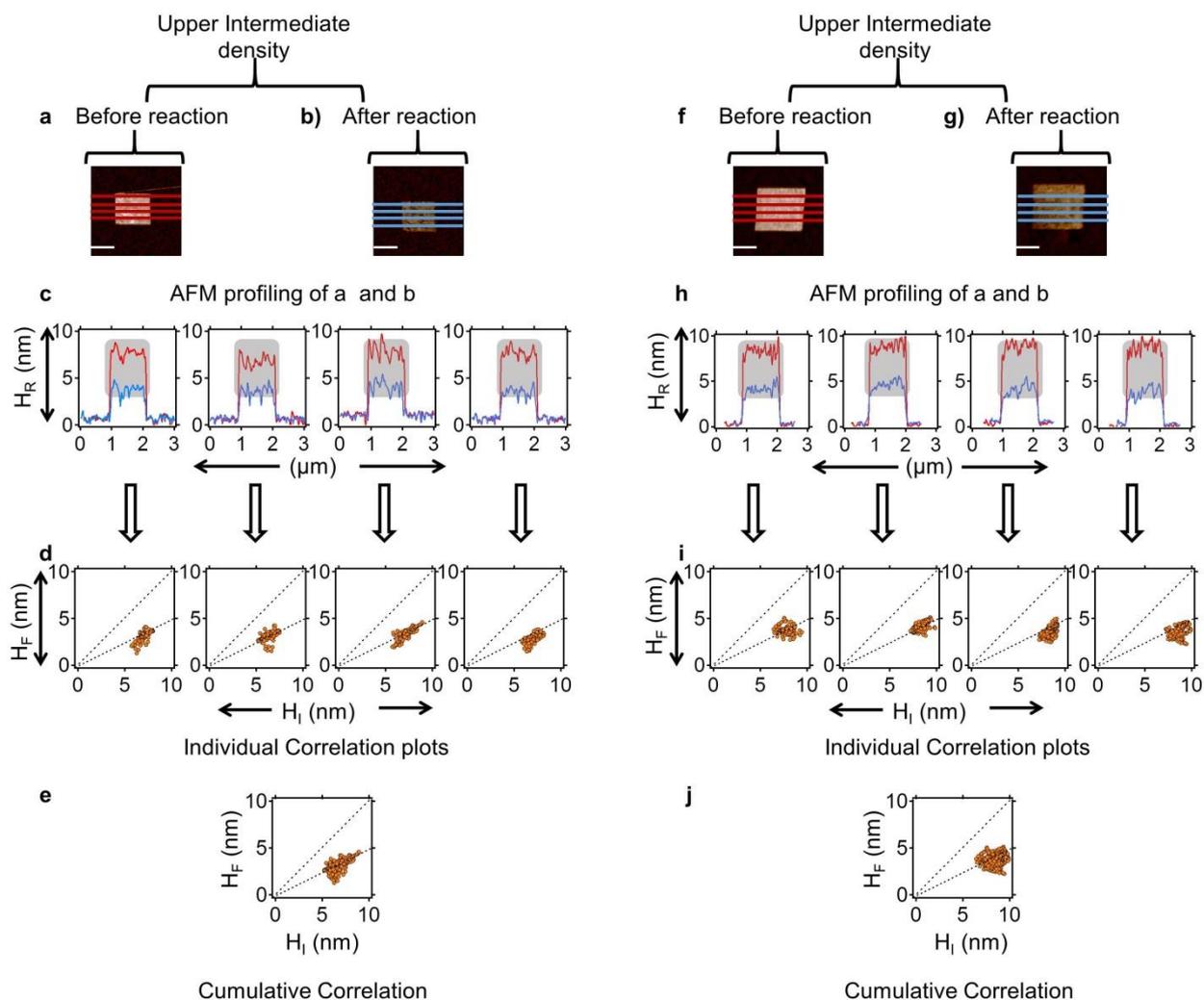


**Figure 2.5:** Side-by-side AFM height profiling of DNA nanoreactors made up of DNA-1 (left) and DNA-3 (right), with intermediate density. For both DNA-1 and DNA-3 nanoreactors, the AFM topographic images (a-b for DNA-1, f-g for DNA-3) were divided into four subunits, with line profile across each subunit. Each line was mapped one-to-one across before and after AFM micrographs. (c, h) Show the one-by-one height profile that corresponds to the one-to-one mapping of the line for before (red) and after (light blue) BamHI action. For all the four height profiles in (c), the relative final height of the DNA-1 molecules is half of the relative initial height. In (h), however, each individual height profile has its unique difference in height. Also, in all of the height profiles, the reduction in height after enzymatic reaction is nearly 25 %. (d, i) represent the collated height waves for each profile in (c, h), reflecting the linear correlation between the initial relative height ( $H_I$ ) and the final relative height of the DNA molecules after the enzymatic reaction. (e, j)

shows the cumulative correlation plot of all the four individual plots in (d, i), with a linear regression fit. Also, the slope of linear fit is represented in form of a column plot. Scale bar is 750 nm.

Furthermore, to quantify the digestion of BamHI within cognate (DNA-1) and non-cognate (DNA-3) nanoreactors, we collected the height waves in each AFM profile in (2.5c, h), with respect to the before and after enzymatic reactions, and translated the waves into peak-to-peak individual correlation plot, which connects the relative initial height waves to the relative final height waves, with two diagonal dashed lines at an angle  $45^\circ$  and  $22.5^\circ$  to the x axis. The dashed line at  $45^\circ$  stands for the state of no digestion, meaning the relative initial height is equal to the relative final height waves i.e.  $H_F = H_I$ , while the dashed line at  $22.5^\circ$  stands for the state of 50 % digestion, i.e. relative final height is equal to half the relative initial height i.e.  $H_F = 0.5H_I$ . With such statistical analysis, we observed that each subunit or subdomain within nanoreactor consists of different density domains, as shown in Fig 2.5d, and i, with each individual plot showing the density width, of relative initial height 5-7 nm with corresponding estimated density 1650-2000 molecules/ $\mu\text{m}^2$ . Despite the heterogeneity in density, BamHI digested completely DNA-1 molecules (cognate molecules), with 50 % reduction in height, as all the height population waves lie on the 50 % dashed line, irrespective of the subdomain density as shown in the Fig. 2.5d. Conversely, in DNA-3 nanoreactor, the action of BamHI was only completed at the low-density domains within the nanoreactor, while the high-density domains remain substantially unmodified by BamHI. This observation is revealed in Fig. 2.5i, as the low-density population (with  $H_I = 5$  nm, estimated density  $\sim 1650$ ) lie on the 50% digestion line and the high-density population tend to be far away from 50% digestion but close to no digestion dashed line, as shown by all the individual plots in (2.5i).

Subsequently, the whole individual correlation plots were collated into a single correlation plot, one for DNA-2 (2.5e) and the other for DNA-3 (2.5j), each cumulative correlation plot is equivalent to the sum of all the individual plots in (2.5d) and (2.5i). Additionally, the cumulative plots were fitted with linear regression line, together with a column plot of linear slope against the relative initial height. Based on the cumulative correlation plots in (2.5e, j), BamHI fully digest DNA-1 molecules regardless of the heterogeneous molecular density, while its action only reach completion in the low-density domains within DNA-3 nanoreactor and the high-density domains remain un-cut.



**Figure 2.6:** Examining the action of BamHI within DNA-1 (left) and DNA-3 (right) molecules at upper intermediate densities. In the first column, (a-b) shows the two AFM micrographs that correspond to the nanoreactor before and after the BamHI action. Each of the nanoreactor was divided into four subunits, with line profile across each subunit, and each line mapped to one another. The same procedure of data analysis was also performed on DNA-3, as shown in (f, g). From the lines profiles, individual AFM height profiles were generated, four height profiles for DNA-1 in (c) and four for the DNA-3 in (h). Each height profile corresponds to the one-to-one mapping (red to light blue of the same subunit) of the line profile across (a, b) and (f, g), respectively. In all of the height profiles in (c) and (h), the relative final height (light blue) of the DNA nanoreactor is half of the relative initial height (red). This is evidently shown in each of the individual correlation plot in (d) and (i), as all the height waves fall on the dashed line of 50% cleavage. (e, j) show the cumulative correlation plot of all the individual plot in (d) and (i) respectively. Scale bar, 750 nm.

In addition to the results in Fig. 2.5, we extended the same approach to quantify the action BamHI in nanoreactors with higher density. We observed the same even action of

BamHI in all the density domains within DNA-1 nanoreactor, as shown in Fig. 2.6 c-d. The estimated densities for DNA-1 molecules within the nanoreactor is 1800-2300 molecules/ $\mu\text{m}^2$ .

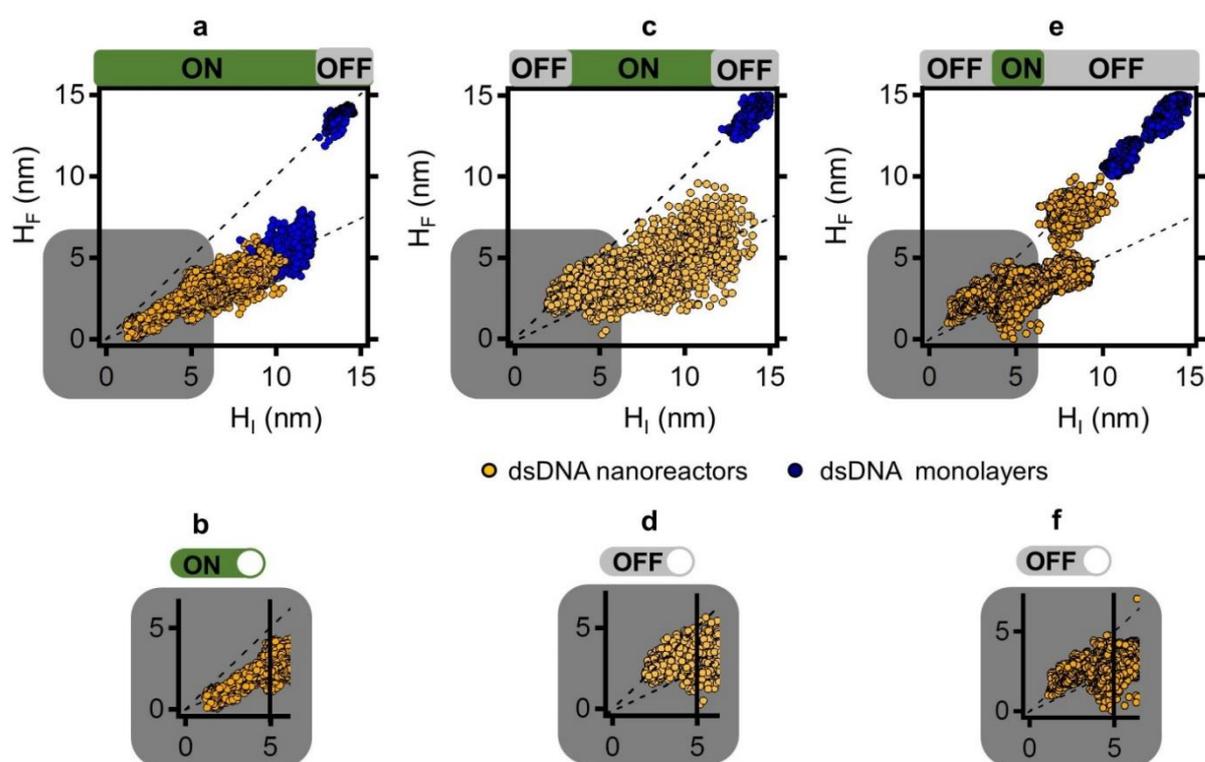
Contrary to the unequal digestion activity associated with BamHI in DNA-3 molecules at the bottom of the intermediate density range shown in (2.5h-i), for the upper part of this range we observed even with equal 50% digestion in all density domains within DNA-3 nanoreactor, as revealed by the AFM height profiles and individual correlation plots in (2.6h-i). The estimated molecular density in this case ranges from 2000-3400 molecules/ $\mu\text{m}^2$ , with relative initial heights of 7-9 nm. This is nearly 1.5-fold higher than the densities in (2.5f). Also, the cumulative correlation plot revealed complete digestion, as the collated height waves corresponding to the relative initial height waves and relative final height waves, lie on the 50% digestion line (dashed line), as shown in Fig. 2.6j.

In summary, there are dual actions of BamHI captured in DNA-3 nanoreactors within the intermediate density regime; with 1650-2000 molecules/ $\mu\text{m}^2$  BamHI has incomplete digestion with low efficiency, above his density frame, i.e. 2000-3400 molecules/ $\mu\text{m}^2$ , BamHI completely digest the DNA-3 molecules with high efficiency, as shown in 2.6j. Evidently, the novel analysis approach allows us to capture the **onset** of the anomalous action of BamHI and its **completion**, as triggered by nanoscale confinement and crowding within the nanoreactors.

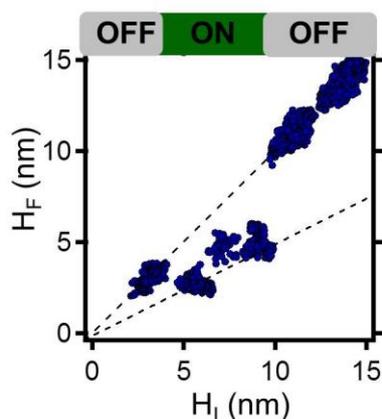
Encouraged by the unprecedented results revealed by the novel side-by-side AFM height profiling approach, we further analysed and quantified the recognitions of BamHI in sequence-dependent nanoreactors, either constructed by nanografting or by spontaneous formation of self-assembled monolayer of dsDNA, within varying molecular densities, ranging from 500 molecules/ $\mu\text{m}^2$  to 5000 molecules/ $\mu\text{m}^2$ . The overall actions of Bam HI are shown in Fig. 2.7 and 2.8.

Remarkably, the overall actions of BamHI enzyme, in relation to sequence-density dependent, are summarized as follows; BamHI cleaves the cognate site (DNA-1) in nanoreactors with height lower than 14 nm (see Fig. 2.7 a). This height value corresponds to a critical DNA density ( $\sim 4500$  molecules/ $\mu\text{m}^2$ ), above which the dsDNA accessibility is negligible due to steric hindrances, this is in agreement with the previous publication of our group<sup>14</sup>. In addition, we observed three distinct and unprecedented behaviours for DNA-2 and DNA-3. For each DNA, we observed three to four distinct regimes. At very low densities (500-1500 molecules/ $\mu\text{m}^2$ ), BamHI cleaves none of the two (see Fig. 2.7d, f). This is in agreement with its known high specificity for the cognate site (DNA-1) in dilute

systems<sup>6</sup>. For nanoreactors with thickness profiles higher than 5 nm (1650-3400 molecules/ $\mu\text{m}^2$ ), BamHI unexpectedly digests both non-cognate sites (DNA-2 and DNA-3). In such nanoreactors, however, the reaction reaches completion only within the lower-density domains, with an exception at the moderately high density ( $\sim$  2000-3400 molecules/ $\mu\text{m}^2$ ), where the BamHI reaction reached completion in DNA-3 nanoreactors (see 2.7e). Also, the action of BamHI reaches full completion within non-laterally confined dsDNA monolayer, made up of DNA-3, only at intermediate densities (1650-3500 molecules/ $\mu\text{m}^2$ ) as shown in Fig. 2.8. Finally, the enzymatic reaction is fully inhibited for nanoreactors with thicknesses higher than 14 nm for DNA-2, and 11 nm for DNA-3 (see Fig 2.7c, e, and 2.8).



**Figure 2.7:** Correlation plots of relative final height versus relative initial height, showing the general sequence-density function of BamHI together with ON-OFF thresholds, indicating the density range where the action of BamHI is ON and the density range where it is OFF. The plots in (a), (c), and (e) highlight the behaviour of Bam HI within density-dependent DNA nanostructures, made up of DNA-1, DNA-2, and DNA-3 respectively. The second row of the Fig., i.e. (b), (d), and (f) clearly shows the notable action of BamHI within low density regime, with respect to DNA-1, DNA-2, and DNA-3 respectively.



**Figure 2.8:** Correlation plots of relative final height versus relative initial height, showing the unexpected action of BamHI within DNA-3 monolayers, revealed by side-by-side AFM height measurements.

## 2.6 Discussion and conclusion

In this study, we have examined the recognition and the specificity of endonuclease within highly confined and crowded DNA compartments. We did so by integrating an AFM nanolithography based method termed nanografting, spontaneous formation of self-assembled monolayers of pre-hybridized dsDNA, with AFM topographic imaging and height measurement. Unequivocally, we demonstrated the effect of nano-scale confinement, compartmentalization, and crowding on the specificity and recognition of BamHI within spatially resolved cognate and non-cognate DNA nanoreactors.

Our unprecedented results show that DNA density is a crucial determinant for the action of BamHI<sup>14,20</sup>. By the aid of novel AFM side-by side profiling analysis, we observed a switch-like action of BamHI within the sequence-dependent DNA nanoreactors, in relation to the heterogeneity in molecular density.

Our interpretation is that the restriction enzyme is physically trapped within nanoreactor with thicknesses higher than 5 nm, as substantiated by previous published work of our group<sup>14</sup>. Thus, the enzyme diffuses within the nanoreactor and the non-canonical reactivity of DNA-2 and DNA-3 is likely an effect of crowding experienced by the enzyme under nanoscale confinement. Previous studies of crowding on biochemical reactions suggest that the DNA molecular density is 1) directly related to enzyme and DNA activities, and 2) inversely related to enzyme mobility within the highly dense DNA phase (diffusion-limited regime)<sup>1</sup>. i.e. high activities and affinities favour the successful reaction, whereas low diffusivity inhibits the enzyme-DNA collisions and thus complex formation. Our results therefore, seem to suggest that under nanoscale confinement the strength of the enzyme-

DNA binding (that varies with the different restriction sites) strongly contributes to balance of such factors, and underlies novel enzyme behaviours that are not observable in solution. In particular, our results indicate the existence of threshold between region in which the enzyme can or cannot work in non-canonical conditions, whereas no study on macromolecular crowding has ever reported such a digital behaviour to our knowledge. In contrast to our systems, the systems investigated in studies on macromolecular crowding are typically comprised of dense, disordered polymers that increase the average volume voidance in solution. The action of an enzyme molecule on its substrate cannot be associate therefore to a well-defined level of crowding. In turn, the effect of crowding on enzyme activity and diffusivity is only an average. Our interpretation is that in our nanoreactors, in which the DNA density is a well-defined spatial dependent variable, enzyme activity and diffusivity assume well defined, position-dependent values, leading to BamHI enzyme behaviours that are distinct over the three DNA sequences employed despite the small differences in sequence, as well as remarkably dependent upon DNA density in a digital fashion. These findings demonstrate that the use of nanoreactors allows studying unprecedented modes of protein-DNA recognition. In general, our results open the door to the possibility of using DNA self-assemblages as nanodevices for investigating biomolecular interactions in complex systems.

## **2.7 Materials and methods**

### **Materials**

Thiol modified (with C6) single stranded DNAs oligomers (ssDNAs) and their corresponding complementary ssDNAs' were purchased from Integrated DNA Technologies (IDT). BamHI-HF (with CutSmart buffer; 50mM potassium acetate ,20mM tris-acetate,10mM magnesium acetate 100µg/ml BSA) was purchased from New England Biolabs, Ipswich, USA .C11 Alkane thiols omega (Top-) substituted with -Oligo-Ethylene Glycol (TOEG 6) (HS-((CH<sub>2</sub>)<sub>11</sub>) -(O-CH<sub>2</sub>-CH<sub>2</sub>)<sub>6</sub> -OH) was purchased from Prochimia Surfaces (Poland), sodium chloride (NaCl), Tris-EDTA, absolute ethanol (99.8%) were all purchased from Sigma Aldrich . All buffer solutions were prepared in ultra-pure water (miliQ-H<sub>2</sub>O), of resistivity 18.2 MΩ·cm at 25 °C and filtered before use with a sterile syringe filter (of 0.22µm pore size).

## DNA-sequences design

Three distinct, 44-bp-long DNA molecules were strategically designed with recognition sites located in the middle. One is provided with the correct BamHI restriction site that ensures specificity for the enzyme (5'-GGATCC-3', termed **DNA-1**), while the other two are provided with sequences that only partially match such restriction site (i.e. 5'-CGATCA-3' and 5'-AGATCA-3', termed **DNA-2** and **DNA-3** respectively). The sequences with their corresponding complementary are listed in Table 2.0.

**Table 2.0:** Thiol modified ssDNA sequences and their complementary

DNA molecules	Sequences
DNA-1	SH-((CH) <sub>2</sub> ) <sub>6</sub> -5'-CAAAACAGCAGCAATCCAAGGATCCGACACCCGATTACAAATGC-3'
DNA-2	SH-((CH) <sub>2</sub> ) <sub>6</sub> -5'-AAAAAAAAAAAA AAAACGATCATGAAGCACATCAGAGTCTCCTAC-3'
DNA-3	SH-((CH) <sub>2</sub> ) <sub>6</sub> -5'-AAAAAAAAAAAAAAAAAAGATCATGAAGCACATCAGAGTCTCCTAC-3'
<b>Complementary sequences</b>	
cDNA-1	5'- GCATTTGTAATCGGGTGTCTGGATCCTTGGATTGCTGCTGTTTTG-3'
cDNA-2	5'-GTAGGAGACTCTGATGTGCTTCATGATCGTTTTTTTTTTTTTTT -3'
cDNA-3	5'- GTAGGAGACTCTGATGTGCTTCATGATCTTTTTTTTTTTTTTTT-3'

## Preparation of ultra-flat gold substrate

Ultra-flat gold substrates were prepared as described in the past publications from our group<sup>14,17</sup>. Briefly, a sequential deposition of gold on freshly cleaved mica was employed using electron beam evaporation in a vacuum chamber (pressure about 10<sup>-6</sup> mbar); firstly, gold was deposited at the rate 0.01 nms<sup>-1</sup> until a film of 20 nm was obtained, then the rate of evaporation was increased to 0.1 nms<sup>-1</sup> until a film of 100 nm thick was formed on freshly cleaved mica (Mica New York Corp., clear ruby muscovite) at room temperature.

The planar gold sheet of (111) crystallographic plane obtained on mica was sliced into few milli-meter squares (approximately 5 x 5 mm<sup>2</sup>) in area using Stanley-199 blade.

To form ultra-flat free gold surfaces, the freshly formed gold surface was first bonded to the polished side of a silicon wafer, with a drop of SU- 8 photoresist adhesive (negative tone photoresist, MicroChem, USA). A sandwiched was obtained, silicon-gold-mica, by pressing the polished section of silicon on the gold section of mica. The silicon-gold-mica sandwich squares were cured at 130°C for at least 24 hours. The samples were then allowed to cool down to room temperature without any external cooling system and this was done to avoid thermal stresses that can result into gold film detaching from the mica substrate.

Without any further surface treatment, the samples were stored at room temperature, ready to be used (after stripping the mica from the gold surface) for self-assembled monolayer preparation.

### **Preparation of top-oligo-ethylene-glycol SAM on ultra-flat gold substrate**

An ultra-flat gold surface was obtained by mechanical stripping the mica substrate from silicon-gold-mica sandwich, followed by immersion and incubation in the solution of 100  $\mu\text{M}$  of top-oligo-ethylene glycol (TOEG, 6,  $(\text{HS}-((\text{CH}_2)_{11})-(\text{O}-\text{CH}_2-\text{CH}_2)_6-\text{OH})$ ) in absolute ethanol (Prochimia and Sigma Aldrich) and 1M NaCl, TE1X (10-mM Tris-HCl, 1-mM EDTA, pH 7.2 at 25°C in Milli-Q water) for 6-hours. This TOEG SAM serves as bio-repellent monolayer for aspecific adsorption of biomolecules on the surface. After the incubation time, serial rinsing was performed in both ethanol and 1M NaCl, TE1X buffer to remove any physically-adsorbed contaminants.

### **Approach 1: Generation of DNA nanoreactors via nanografting**

Nanografting experiments were performed as follows; A TOEG SAM passivated gold substrate was adhered firmly on the home-made liquid cell. Next, a ssDNA solution (0.1  $\mu\text{M}$ -1  $\mu\text{M}$  ssDNA, 1M NaCl, TE1X) was evenly dispensed on the SAM and it was transferred onto the AFM X-Y scanner stage. Preliminary imaging was performed in tapping mode at low force to obtain the topographic image of the TOEGSAM surface and also to select a flat and clean section for nanografting. Afterwards, a 1 x 1  $\mu\text{m}^2$  section was selected for the grafting process and concomitantly at high force ( $\approx 120\text{nN}$  in contact mode); TOEG 6 molecules within the selected 1 x 1  $\mu\text{m}^2$  were substituted with thiol-modified ssDNA molecules in solution. However, to obtain several patches with different molecular densities, grafting parameters such as patch dimension, the number of times AFM tip over-writes a selected area during nanografting, concentration of ssDNA in solution, and applied nominal force, were tuned during the nanografting process. After the grafting process, a 20 X 20  $\mu\text{m}^2$  section that contained the grafted patches was scanned at very low force (high set point) in tapping mode. Domain reference (the section that contains the grafted patches) was saved using SV-align software<sup>21</sup>.

### **Hybridization of thiol-modified ssDNA nanoreactors**

After nanografting, the sample was thoroughly rinsed with 1M NaCl, TE1X buffer (DNA-free buffer). For hybridization, 1  $\mu\text{M}$  complementary solution was evenly dispensed on to

the surface and the hybridization takes place at 37°C for 1-hour, using Binder oven (model FD 115, BINDER GmbH, tuttlingen, Germany). After hybridization, the sample was copiously rinsed with 1M NaCl, TE1X buffer (DNA-free buffer). The aforementioned procedures are applicable to DNA-1, 2 and 3 nanoreactors.

### **AFM nanolithography and imaging**

All AFM experiments (scanning, nanolithography, height detection, and recognition of restriction enzymatic reaction) were carried out with an MFP-3D Classic (also called stand-alone) and MFP-3D-BIO AFM, (Asylum research, Santa Barbara, CA).

For nanografting, commercially available pyramidal silicon etched probes, NSC19/no Al and NSC 18 /no Al of spring constant 0.6N/m and 2.3 N/m respectively (Mikro-Masch, Germany) were used.

For imaging the nanoreactors before and after restriction enzymatic reactions, a soft probe (CSC 38 /no Al, spring constant 0.03N/m, Mikro-Masch, Germany) was used.

### **Approach 2: Self-assembled monolayer formation of double stranded DNA**

To obtain self-assembled monolayer formation of dsDNA on an ultra-flat gold substrate, double stranded DNA SAMs were formed by incubating equal molar concentration of both thiol-modified and complementary ssDNA in the presence of 100mM MgCl<sub>2</sub> and 1M NaCl, TE1X buffer (10-mM Tris-HCl, 1-mM EDTA, pH 7.2 at 25°C in Milli-Q water) at 37°C for 1-hour. This step is called pre-hybridization. After the pre-hybridization step, absolute ethanol (of a defined volume) was added to the pre-hybridized DNA solution, followed by ultra-flat gold substrate for the SAM formation. At this point, it is important to state that different molecular densities were obtained by tuning the concentration of both thiol-modified and complementary ssDNA, the incubation time for SAM formation, and the volume % of ethanol. After elapsing of the SAM formation time, the sample was rinsed twice in DNA-free 1M NaCl, TE1X buffer (10-mM Tris-HCl, 1-mM EDTA, pH 7.2 at 25°C in Milli-Q water), then passivated with 100µM TOEG6 in 1M NaCl, TE1X buffer for 15 minutes and finally rinsed twice again in 1M NaCl, TE1X buffer after the passivation step. The obtained dsDNA SAM on gold substrate was fixed within the home-made liquid cell, covered with DNA-free 1M NaCl, TE1X buffer and it was transferred onto AFM X-Y scanner stage for imaging and experimentation.

Preliminary imaging was performed in tapping mode at low force to obtain the topographic image of the dsDNA SAM surface and also to select a flat and clean section for nanografting.

Nanografting of TOEG6 was performed within the dsDNA SAM in 10- $\mu$ M TOEG6, 1-M NaCl, and TE1X buffer. This is a typical example of negative nanografting. The nanografting was performed in contact mode using the following parameters; applied load  $\approx$  120nN (high set point), point and lines: 256, scan rate: 1Hz and scan angle 90°, and cantilever probe: NSC 19/ no Al.

### **AFM measurements**

To carry out an experiment it is necessary to perform tip-sample alignment using the SV-alignment software. SV-alignment software simply superimposes the reference tip-sample image on the video panel and the live tip-sample image. This alleviates the problem of loss of experimental domain, where the grafted patches were located on the sample, and support continuous experiment on the same experimental domain.

MFP-3D Classic and MFP-3D bio were both used to characterize and measure the topographic height of the dsDNA patches before and after the enzymatic reaction. The characterization and the measurement were taken in contact mode with soft probe (CSC38/ no Al) at the lowest force ( $\approx$  0.2nN) within the DNA free buffer (1M NaCl, TE1X buffer).

### **Restriction enzymatic reaction**

BamHI-HF (from *Bacillus amyloliquefaciens*, type II restriction enzyme), is the restriction enzyme used in this experiment. As a follow-up procedure after the measurement of the topographic height of the DNA nanoreactors, the sample was rinsed with 1M NaCl, TE1X buffer. Hence to equilibrate the macro-environment for restriction enzymatic reaction, the sample was rinsed and incubated with enzymatic reaction buffer, cutsmart buffer (1X concentration, 50mM Potassium acetate, 20mM Tris acetate, 10mM Magnesium acetate, 100 $\mu$ g/ml BSA) for 10 minutes.

The restriction reaction was carried out at 0.2 U/ $\mu$ l ( $\approx$  20 nM) at 37°C for 1-hour. At the completion of the reaction, the sample was rinsed with free enzyme cutsmart buffer and subsequently rinsed with DNA-free 1M-NaCl, TE1X buffer.

## 2.8 References

1. Elli, R. J. Macromolecular crowding: obvious but underappreciated. **4**, 597–604 (2001).
2. Pingoud, a & Jeltsch, a. Recognition and cleavage of DNA by type-II restriction endonucleases. *Eur. J. Biochem.* **246**, 1–22 (1997).
3. Vasu, K. & Nagaraja, V. Diverse functions of restriction-modification systems in addition to cellular defense. *Microbiol. Mol. Biol. Rev.* **77**, 53–72 (2013).
4. Types of Restriction Endonucleases | NEB. at <<https://www.neb.com/products/restriction-endonucleases/restriction-endonucleases/types-of-restriction-endonucleases>>
5. Chan, S. & Biolabs, N. E. Restriction Endonucleases : Molecular Cloning and Beyond. 1–6 (1965).
6. Viadiu, H. & Aggarwal, A. K. Structure of BamHI bound to nonspecific DNA: A model for DNA sliding. *Mol. Cell* **5**, 889–895 (2000).
7. Sun, J., Viadiu, H., Aggarwal, A. K. & Weinstein, H. Energetic and Structural Considerations for the Mechanism of Protein Sliding along DNA in the Nonspecific BamHI-DNA Complex. *Biophys. J.* **84**, 3317–3325 (2003).
8. Paillard, G. & Lavery, R. Analyzing Protein-DNA Recognition Mechanisms. *Structure* **12**, 113–122 (2004).
9. Morozov, A. V., Havranek, J. J., Baker, D. & Siggia, E. D. Protein-DNA binding specificity predictions with structural models. *Nucleic Acids Res.* **33**, 5781–5798 (2005).
10. Selvaraj, S., Kono, H. & Sarai, A. Specificity of protein-DNA recognition revealed by structure-based potentials: Symmetric/asymmetric and cognate/non-cognate binding. *J. Mol. Biol.* **322**, 907–915 (2002).
11. Rosenberg, J. M. Structure and function of restriction endonucleases. *Curr. Opin. Struct. Biol.* **1**, 104–113 (1991).
12. David T F Dryden. Fig. - Nature Structural & Molecular Biology. at <[http://www.nature.com/nsmb/journal/v11/n9/fig\\_tab/nsmb0904-804\\_F1.html](http://www.nature.com/nsmb/journal/v11/n9/fig_tab/nsmb0904-804_F1.html)>
13. Goodsell, D. S. Restriction Enzymes. *RCSB Protein Data Bank* (2000). doi:10.2210/rcsb\_pdb/mom\_2000\_8
14. Castronovo, M. *et al.* Two-dimensional enzyme diffusion in laterally confined DNA monolayers. *Nat. Commun.* **2**, 297 (2011).

15. Liu, M., Amro, N. A., Chow, C. S. & Liu, G. Y. Production of Nanostructures of DNA on Surfaces. *Nano Lett.* **2**, 863–867 (2002).
16. Liu, M. & Liu, G. Y. Hybridization with nanostructures of single-stranded DNA. *Langmuir* **21**, 1972–1978 (2005).
17. Mirmomtaz, E. *et al.* Quantitative study of the effect of coverage on the hybridization efficiency of surface-bound DNA nanostructures. *Nano Lett.* **8**, 4134–4139 (2008).
18. Peled, D., Daube, S. S. & Naaman, R. Selective enzymatic labeling to detect packing-induced denaturation of double-stranded DNA at interfaces. *Langmuir* **24**, 11842–11846 (2008).
19. Bosco, A. *et al.* Hybridization in nanostructured DNA monolayers probed by AFM: theory versus experiment. *Nanoscale* **4**, 1734 (2012).
20. Matteo Castronovo,<sup>†,‡,§,||,∇</sup> Slobodanka Radovic,<sup>⊥</sup> Christian Grunwald,<sup>†,‡,¶</sup> Loredana Casalis,<sup>†,§</sup> Michele Morgante,<sup>⊥</sup> and G. S. Control of steric hinderances on restriction enzymes reactions with surface-bound DNA nanostructures. *Nano Lett.* (2008).
21. Panel, T. A. Sample Alignment Software Tool SVAAlign . ipf.

### 3.0 Results and discussion on AFM-based characterization and determination of binding affinity of *in silico* optimized peptides towards a biomarker-protein\*

*In this chapter, all the experimental validation based on atomic force microscopy, to determine the binding affinity to  $\beta 2m$  of peptides computationally designed by Dr Sara Fortuna and Dr Miguel Soler, were performed. I set up the AFM-based methods, which includes nanografting, DNA-directed immobilization, differential height measurements, and its analysis using the Igor pro software.*

Here we show that it is possible to target a surface exposed binding site by designing *de novo* binder able to immobilize its target protein on a surface. To this aim, our collaborator (Dr Miguel Soler and Dr Sara Fortuna, MONALISA Lab, University of Udine) applied a stochastic evolutionary algorithm<sup>1-3</sup>. This algorithm is capable of generating peptides suited for capturing a protein in solution reaching the same micro-molar affinity as that obtained by selecting peptides either by phage display library, or by natural evolution.

To push forward the binding affinity of the optimized probes towards their target and address some of the drawbacks present in previous design strategies<sup>1,3</sup>, the theoretical group has optimized, and tested a version of the algorithm which enhances the sampling of the bound state with molecular dynamics (MD) based techniques<sup>2</sup>.

As a test model, we have chosen  $\beta 2$ -microglobulin ( $\beta 2m$ ), a well characterized<sup>4,5</sup> 12kDa protein included in the OVA 1 test panel for the prediction of ovarian cancer aggressiveness<sup>6</sup>.  $\beta 2m$  also accumulates in patients undergoing hemodialysis and eventually leading to amyloidosis<sup>7</sup>. Its three-dimensional structure was first reported by Becker *et al.*<sup>8</sup> in 1985 using X-ray crystallography techniques, and its size was reported to be 4.5 x 2.5 x 2.0 nm<sup>3</sup> in volume, as depicted in Fig. 3.0a. As a binding site, they selected the region known to bind the human histocompatibility antigen<sup>5</sup>. The selected binding site is exposed to the solvent and easily accessible to the designed peptides (see Fig. 3.0a). For that binding site, cyclic peptides were designed and cyclized by a single disulphide bridge<sup>9</sup>. The cyclic

---

\* Part of this chapter was recently published in form of an article in Physical Chemistry Chemical Physics, DOI: 10.1039/C6CP07807A.

structure is expected to reduce the conformational freedom of the unbound peptide reducing unfavourable entropic contributions to the complex formation<sup>10</sup>.

Subsequently, we developed an experimental-based validation approach that combines nanografting, DNA-directed immobilization and atomic force height measurements, to immobilize the chosen cyclic peptide, pep461, on an ultra-flat gold substrate, to form DNA-peptide arrays, which can bind to the solvent-exposed site on the Beta-2-microglobulin ( $\beta$ 2m). Lastly, with respect to the synergy approach in the framework of AFM, we will show that pep461 has a micro-molar binding affinity towards  $\beta$ 2m, of 5-fold of magnitude higher than that of the binding affinity obtained via SPR technique (see Appendix B for detailed results on SPR experiments).

### 3.1 Computer-based design, generation and screening of cyclic peptides

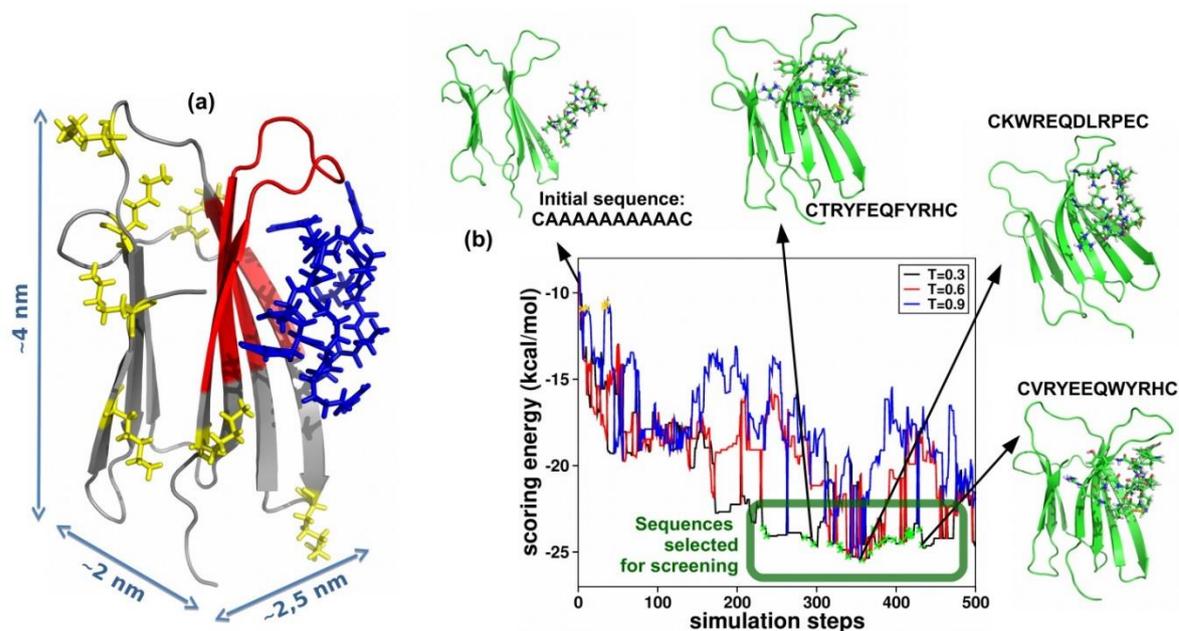
The design approach is based on a combination of replica exchange molecular dynamics (REMD)<sup>11</sup>, cluster analysis<sup>12</sup>, and replica exchange Monte Carlo<sup>13</sup>. The algorithm has been implemented in a bash script embedding functionalities from GROMACS<sup>14</sup>, AMBERTOOLS<sup>15</sup> and AutoDock Vina<sup>16</sup>. At each step (i) a mutation of the sequence is attempted and the system is minimized, (ii) a replica exchange molecular dynamics (REMD) simulation is performed to sample the binder/target conformational space, (iii) representative conformations are selected by clustering the obtained configurations with the Daura's algorithm<sup>12</sup>, and (iv) the conformation with the lowest AutoDock Vina binding score (from now on scoring energy) is chosen and the attempted mutation is accepted or rejected following the Metropolis criteria:

$$P_{\text{acc}} = \min[1, \exp[-(E_{\text{new}} - E_{\text{old}})/T]] \quad (\text{i})$$

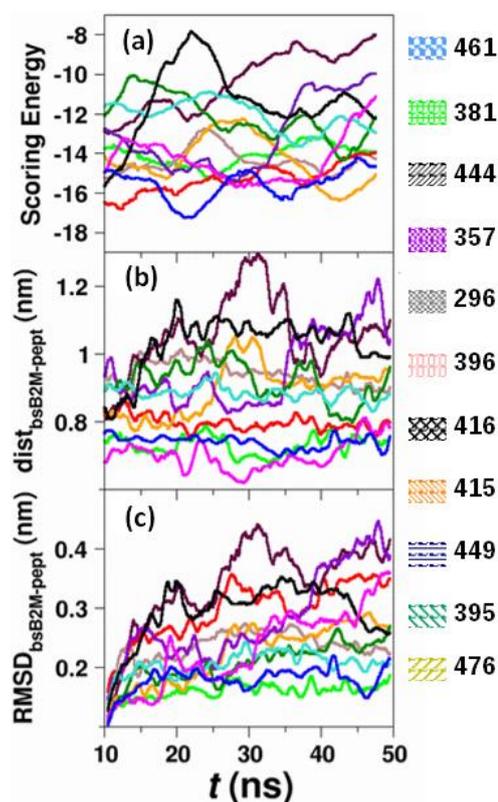
where  $E_{\text{old}}$  is the scoring energy of the old configuration  $E_{\text{new}}$  is that of the attempted mutation, and  $T$  is here a parameter for tuning the Metropolis acceptance probability  $P_{\text{acc}}$ .

Our theoretical collaborator run the optimization algorithm twice, in both instances by starting from the same polyalanine/ $\beta$ 2m configuration (CAAAAAAAAAAC peptide in the inset of Fig. 3.0b). In both cases, the optimization was run in parallel at three values of  $T$  (0.3, 0.6 and 0.9 kcal/mol). Exchanges among the replicas were attempted at every step by following a standard parallel tempering scheme<sup>13</sup>. A typical optimization run is shown in Fig. 3.0b where the evolution of the scoring energy between the  $\beta$ 2m and the peptide is

plotted along the mutation number. The score appears to converge after 200 mutations where the scoring energy settles to values between -24 and -25 kcal/mol. Note the scoring energy is subjected to (possibly large) errors<sup>3</sup> in its absolute value and that it is only in a relative way. In the subsequent production phase, that lasts about 300 mutations, the scoring energy does not change significantly, and the peptides stay well in contact with the protein surface. Among the sequences explored after 200 mutations we selected 20 peptides for further computational screening (see insets of Fig. 3.0b).



**Figure 3.0:** (a) Structure of  $\beta$ 2-microglobulin (gray) bound to the optimized peptide 461 (blue). The  $\beta$ 2m binding site is composed by A, B, E, D  $\beta$ -sheet-forming strands (red). The  $\beta$ 2m lysine residues are also highlighted (yellow). (b) Evolution of the scoring energies of the three Monte Carlo replicas ( $T=0.3, 0.6, 0.9$ ) during a typical run. Low score and soluble peptides (green crosses) are selected for successive screening. Selected configurations along the optimization path are shown in the insets.



**Figure 3.1:** Primary screening: MD simulation results. (a) Average scoring energy, (b) Separation distance between  $\beta$ 2m binding site (BS- $\beta$ 2m) and peptide, and (c) RMSD of BS- $\beta$ 2m+peptide along the simulation time.

Furthermore, the solubility of the computationally optimized peptides was assessed with the Innovagen's peptide calculator webtool<sup>17</sup>. Then, among the peptides predicted to have “Good water solubility”, 20 peptides were chosen with the lowest scoring energy towards the  $\beta$ 2m (Fig. 3.0b). Their sequences are shown in Table B.1. Next, 50 ns MD simulations of each peptide/  $\beta$ 2m complex in water to identify the optimum candidate binders, followed by screening based on a number of observables: the number of hydrogen bonds between peptide and  $\beta$ 2m, the solvent accessible surface area (SASA) of the hydrophobic groups of the binding site of  $\beta$ 2m and of the peptide, and the average root mean square fluctuations (RMSF) of peptide backbone showed negligible differences among candidates, while the scoring energy, the distance between each peptide and the  $\beta$ 2m binding site, and the root mean square deviation (RMSD) between the peptide and the  $\beta$ 2m binding site were a valuable guide in selecting good candidates for experimental screening ( see appendix B for Fig. B6). In particular, for the experimental secondary screening, peptides with both the lowest average values and a stable evolution of these descriptors along the simulated time were selected.

In more detail, as the *scoring* energy of the peptides towards the  $\beta$ 2m was distributed between -20 and 0, peptides exhibiting energies lower than -8 kcal/mol, or with energy fluctuations smaller than 4 kcal/mol were kept (Fig. 3.1a). Peptides with a separation distance from their centre of mass to that of their binding site greater than 1.2 nm were excluded, keeping those with distances between 0.6 nm and 1.2 nm (Fig. 3.1b). Finally, as the RMSD indicates whether either the peptide or the binding site change their conformation (Fig. 3.1c), peptides with large conformational changes (more than 0.5 nm) were discarded, as well as those whose RMSD or separation distance was increasingly diverging along the simulation time. Moreover, peptides differing by only one residue of similar chemical properties, such as Thr/Ser, were also discarded. This primary screening led to eleven peptide sequences (highlighted in Table B2 of the appendix B).

At the end of computation screening, a total number of eleven peptides were subjected to successive experimental screening by SPR, and six out of eleven peptides gave dose response signals (see Table B1 in appendix B). SPR confirmed them to bind to  $\beta$ 2m with  $\mu$ M binding affinity (see Fig. B2). From the six shortlisted candidates, pep461 was chosen among the three peptides with a linear dose response, due to its highest signal/noise response and does not contain histidine (His) residue, thus reducing possible problems of pH control during long MD simulations. Also, ITC confirmed the binding (see Fig. B3), and NMR together with further simulations confirmed the binding site (see Fig. B5). Subsequently, AFM-based experimental validation was developed to quantitatively characterize in a label-free way, the affinity of pep461 towards solvent-exposed binding site of  $\beta$ 2m.

### **3.2 AFM-based experimental validation of cyclic peptide461 towards $\beta$ 2m**

In parallel with the SPR, NMR, and ITC characterization carried out by our collaborators at the University of Naples Federico II and University of Udine respectively (see the detailed experimental methods and results reported in Appendix B), we setup an experimental validation approach, which is based on atomic force microscopy imaging, nanolithography, DNA-directed immobilization, and topography differential height measurements. This approach enables us to quantitatively characterize the binding affinity of chosen cyclic decapeptide, i.e. pep461, towards  $\beta$ 2m in a label-free way.

### 3.2.1 Generation of peptide arrays via nanografting and DNA-directed immobilization

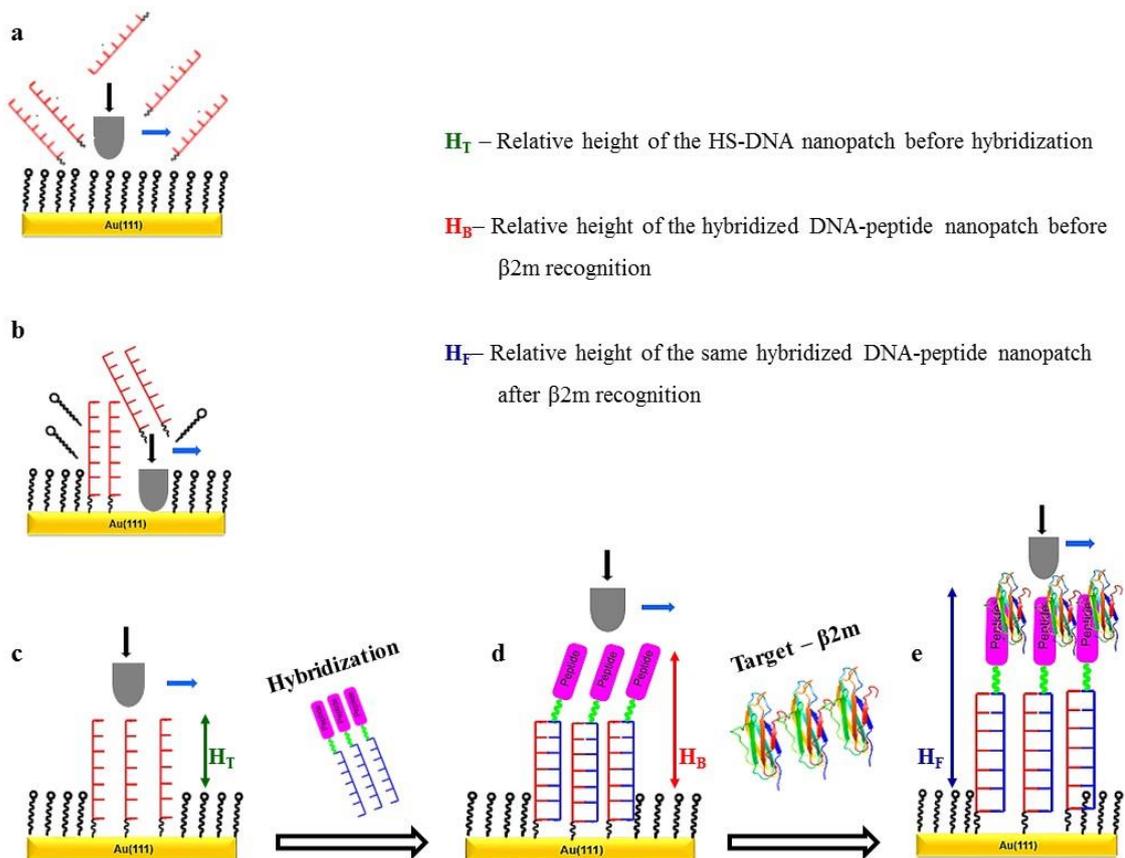
Our strategy to immobilize the peptide on an ultra-flat gold surface through the combination of nanografting and DNA-directed immobilization (DDI), was motivated by their fruitful applications in generating antibody arrays on an ultra-flat gold surface<sup>18,19</sup> and aptamer-based arrays<sup>20</sup>. However, to obtain a bioactive peptide arrays that result from the transformation of DNA arrays via hybridization with the complementary DNA-peptide conjugates, it becomes important for us to use: (1) highly robust conjugation chemistry to create DNA-peptide conjugates, and (2) generate DNA arrays with optimum surface density that does not hinders the functionality of the arrays in the succeeding biochemical reactions. Starting with the conjugation process, we employed NHS-ester-amine conjugation chemistry to conjugate pep461 with amine-modified ssDNA (with C6 chain at the 5'), using the Solulink conjugation kit<sup>21</sup> (see methods section in chapter 4 for details). At the end of the conjugation process, we quantified and determined the concentration of DNA-peptide conjugate using UV-Vis spectrometer.

### 3.2.2 Optimization of the surface density of DNA arrays

To determine the optimum surface density of thiol-modified DNA arrays, we employed nanografting<sup>22,23</sup> (see Fig. 3.2a-c) to generate DNA arrays of different surface densities. Nanografting as discussed in chapter 1 and 2, is a tip-induced self-assembling process and a member of the AFM-based nanolithography technique, which enables the immobilization of biomolecules within the monolayer of bio-repellent alkyl thiol on the surface at a relatively high applied load (~120 nN). Additionally, it also provides the tools to finely-control the density of biomolecules during nanografting by; (1) varying the concentration of the biomolecule solution, in our case it is a solution of thiol-modified DNA molecules; (2) the applied load and (3) the number of times the AFM tip over-writes the selected area<sup>22,23</sup>. In our experimental section, we kept the applied load (~ 120 nN) and the concentration of the thiol-modified ssDNA solution (1  $\mu$ M) constant, while we varied the number of times that the AFM tip over-writes the selected area (1 x 1  $\mu$ m<sup>2</sup>) during nanografting process. Following the calibration scale for surface coverage, tagged S/A, which was developed by Scoles and co-workers<sup>23</sup>, and it is express as;

$$S/A = \frac{R \times N}{W} \quad (\text{ii})$$

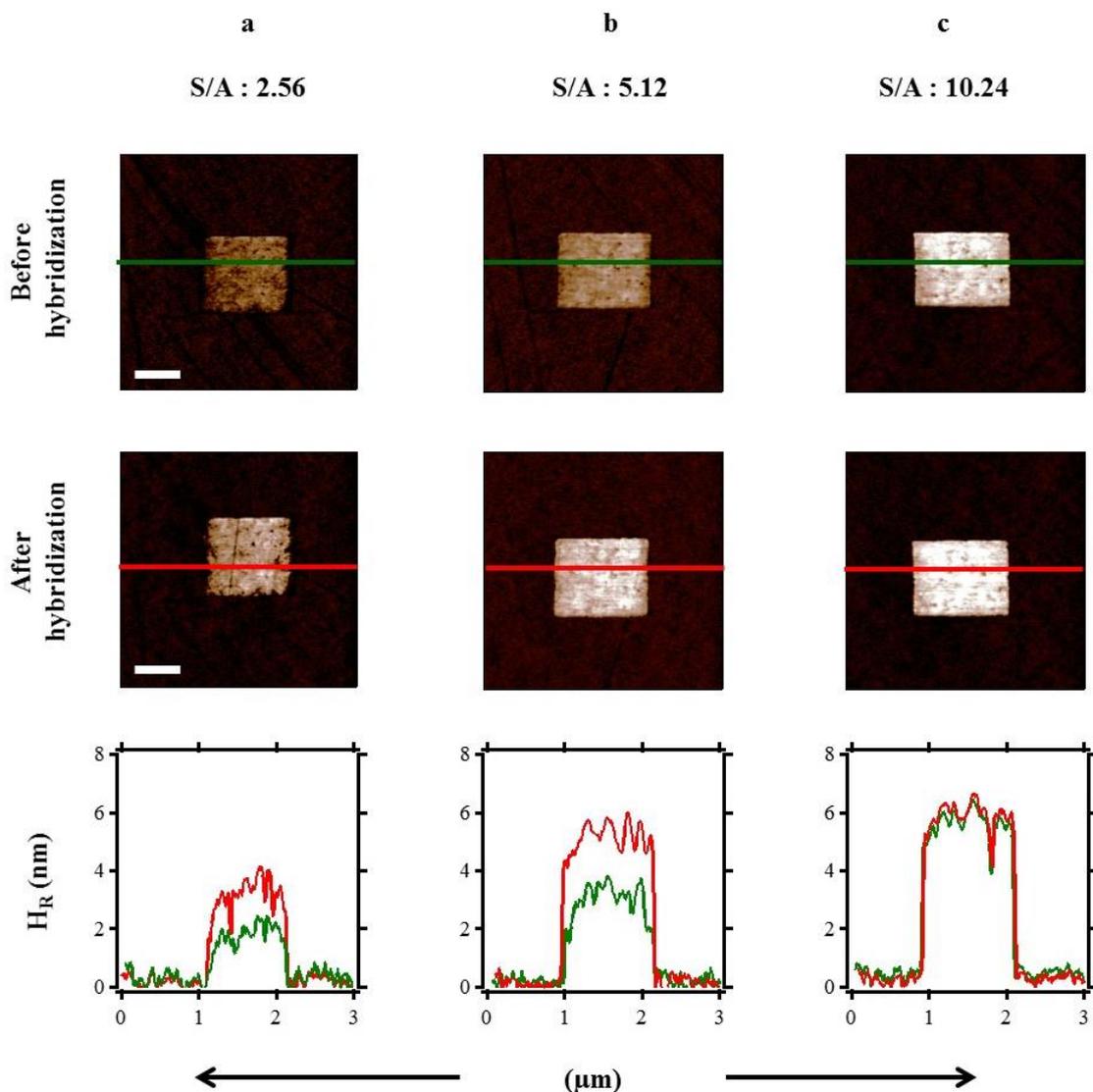
where  $R$  is the radius of curvature of the AFM tip in contact with surface during nanografting process,  $N$  is the number of times the AFM tip over-writes the selected area, and  $W$  is the width of the nanopatch in  $\mu\text{m}$ . In our experimental section,  $R$  and  $W$  were kept constant since we used only one type of AFM cantilever, i.e. NSC 18 / no Al, which has  $\sim 10$  nm as radius of curvature in contact with the surface, and all the patches generated have width of  $1 \mu\text{m}$ . The only parameter that was varied is  $N$ . The master panel of AFM workspace allows us to vary  $N$ , which is the pixel lines, from 96 to 1536, as the AFM tip moves from one addressable area to another on the bio-repellent surface during the nanografting section. By inserting all the parameters into equation (ii) above, we obtained  $S/A$  numbers that range from 1 to 15. With such tool, we generated DNA nanoarrays of varying surface coverage ( $S/A$ ), followed by the acquisition of AFM topographic images, then hybridization with a known concentration of DNA-peptide conjugates at  $37^\circ\text{C}$  for 1 hour. Next, we obtained AFM topographic images of the same DNA arrays, which are now hybridized DNA-peptide arrays, with inherent varying surface density and we then measured the height profiles before (green colour) and after (red colour) the hybridization as depicted schematically in Fig. 3.2c-d.



**Figure 3.2:** AFM-based approach for immobilization of cyclic peptide and detection of  $\beta$ 2m-cyclic peptide interactions on surfaces. The schematics represent the step-by-step procedures of the AFM nanolithography approach termed Nanografting; (a) Scan of the AFM tip at a low nominal load ( $\approx 0.2$  nN) over a given section of a monolayer of protein-repellent ethylene glycol-terminated alkyl thiol on an ultra-flat gold surface (with (111) crystallographic plane), embedded in the  $1 \mu\text{M}$  solution of thiol-modified ssDNA molecules, followed by the removal of the alkyl thiols from the pre-selected section of the ultra-flat gold substrate at high nominal force ( $\approx 120$  nN) and grafting of C6-thiol modified single stranded (ss)DNA in place of the previous alkyl thiol on the surface as shown in (b). This yields laterally confined ssDNA molecules with respect to neighbouring monolayer of protein-repellent alkyl thiol as represented in (c). Subsequently, the complementary DNA-peptide conjugates were immobilized onto (c) by means of DNA-directed immobilization (DDI) through sequence-specific base pairing, and resulted into laterally confined DNA-pep461 as shown in (d). After the successful generation of (d), the DNA-pep461 assemblage is ready to be used for protein recognition (e), in our case  $\beta$ 2-microglobulin ( $\beta$ 2m).

For example, in Fig. 3.3, in the first row, we have the products of nanografting process, presented as AFM micrographs of DNA nanopatch surrounded with the bio-repellent alkyl thiol, which is ethylene-glycol (EG<sub>6</sub>) terminated alkyl thiol, from the lowest surface density to the highest surface density, before the hybridization process. Qualitatively, from the left to the right, the brightness of the nanopatch within the surrounding bio-repellent alkylthiol increases as the S/A increases from 2.56 to 10.24, i.e. the grafted DNA molecules increases within selected area as a function of the number of times the AFM tip over-writes the selected area. In the second row, we have the resulted AFM micrographs of the DNA molecules in the first row after the hybridization with the 500-nM solution of complementary DNA-peptide conjugates. Obviously, the brightness increases with the same trend and if we compare each nanopatch with respect to its prior image before the hybridization, say for example the upper and lower micrograph in Fig. 3.3a, we can clearly see that the hybridization was successful. Beyond the qualitative results, we employed AFM height measurements to compare the relative heights of DNA arrays to the peptide arrays, i.e. relative height profiles of the AFM micrographs before and after hybridization. Notably, the relative height profiles ( $H_R$ ) in the third row of Fig. 3.3 clearly show that the height of the DNA arrays (green colour), increases as the surface density (S/A) increases with an incremental step from left to the right-hand side of the Figure. Likewise, the relative height profiles after the hybridization step (red colour) increases with respect to  $S/A = 2.56 - 5.12$ , but at the very high density ( $S/A = 10.24$ ), the relative height profile

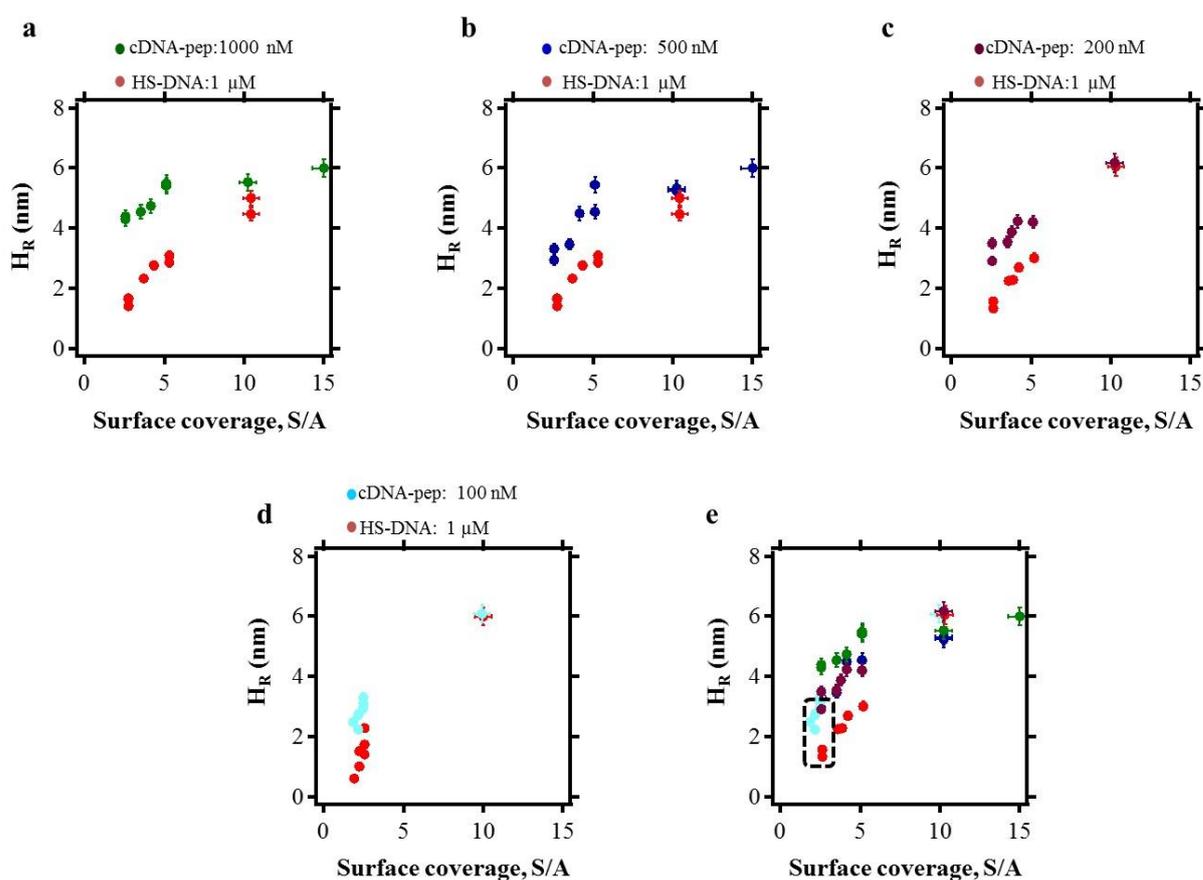
corresponding to before and after hybridization is equivalent to one another. These results agree with the results obtained by Scoles and co-workers<sup>23</sup>, even though the complementary DNA molecules used for hybridization by authors were not conjugated with any protein or peptide.



**Figure 3.3:** Shows an array of nanoassemblies, each generated with different surface coverage (S/A), from the least (left column) to the highest (far right). The first row shows the AFM topographic micrographs obtained after the immobilization process of the thiol-modified ssDNA nanoassemblies through nanografting, before hybridization process. The same AFM micrographs are shown in the second row after the hybridization process with the joining of DNA-pep conjugates of 500-nM concentration. Lastly, the third row shows the AFM height profiles that correspond to the line profiles across the AFM micrographs in the first row and the second row; green line across the AFM micrographs of density dependent thiol-modified ssDNA nanopatches before

hybridization; red line across the same AFM micrographs after the hybridization with complementary DNA-pep conjugates of 500-nM concentration. Scale bar; 500 nm.

To observe the overall effect of surface density of DNA arrays on the hybridization, we performed additional experiments using the same strategies; generating thiol-modified ssDNA arrays of varying surface density, followed by hybridization with complementary DNA-peptide conjugate of different concentration, ranging from 100 nM-1000 nM, and later compare the relative heights before after hybridization, as measured by AFM height measurements.



**Figure 3.4:** AFM height measurements as a function of surface coverage. Each of the plot represent the relationship between the measured relative heights of the nanografted thiol-modified ssDNA (HS-DNA) arrays before and after the hybridization with the complementary DNA-peptide conjugate molecules at different concentration of the conjugate; (a) 1000-nM, (b) 500-nM, (c) 200-nM, (d) 100-nM, and (e) the cumulative plot of a-d, against the surface coverage, tagged S/A. The optimum density selected, S/A = 2.56, is highlighted with the rounded-dash cylindrical shape in (e).

The plots in Fig. 3.4 show the results obtained in several independent experimental sections with respect to different concentrations of DNA-peptide conjugates used during the

hybridization. From the figure, we can deduce that the relative heights of the DNA arrays with respect to the surrounding bio-repellent alkyl thiol increases exponentially with respect to the surface coverage ( $S/A$ ). Also, after the process the relative heights increase with the same trend, except at saturation, where the relative heights before the hybridization are equal to relative heights after the hybridization.

In general, the relative heights of the DNA arrays increase with respect to the concentration of the conjugates used for hybridization, only within the value of  $S/A = 1-6$ , afterwards the relative heights reach a plateau, where the relative heights before the hybridization (red dots) is equal to the relative heights after the hybridization (cyan, blue, green dots), with respect to the concentration of the DNA-peptide conjugates for a given experimental event (see Fig. 3.4 e). Based on the above observations, we can infer that the hybridization is effective at the low range of  $S/A$ , meaning at low density, while at high density it becomes difficult to characterize the hybridization via AFM topographic height measurements. However, the previous studies from our group<sup>23,24</sup> showed that the hybridization efficiency at very high density is  $\sim 50\%$ , as revealed by AFM compressibility measurements and theoretical modelling.

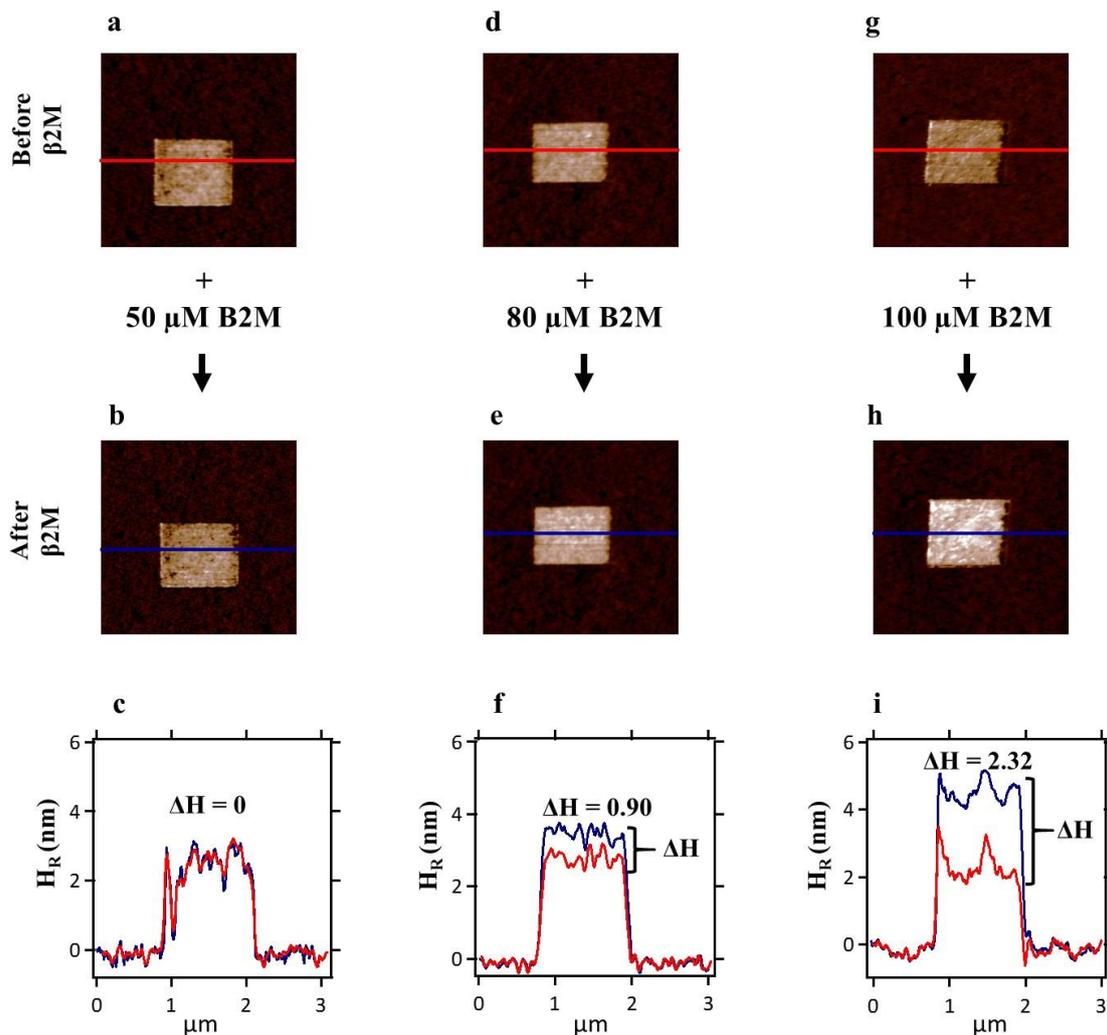
For our experimental validation setup, it is logical to work with an optimum density that is far from saturation plateau of Fig. 3.4 (e), to avoid inaccessibility of the  $\beta 2m$  to the peptide arrays due to steric hindrances and polymerization of the peptides. As such we chose surface density ( $S/A$ ) of 2.54 as the optimum density, with which we can carry-out hybridization, followed by protein recognition without any bias from the surface coverage of the DNA arrays. Also, here we used 100 nM concentration of DNA-peptide conjugates for the hybridization while we used 200 nM in chapter 4. Our choice of concentrations was based on the DNA-peptide conjugates yield obtained after conjugation.

### **3.3. Functionality of DNA-pep461 nanoassemblages and determination of the dissociation constant by Atomic Force Microscopy (AFM)**

Now that we have identified the optimum density for the DNA arrays, the next goal is to determine the binding affinity of DNA-pep461 arrays towards  $\beta 2m$ , to achieve this, we generated three DNA-pep461 assemblages confined within protein-repellent ethylene terminated alkylthiol monolayer on an ultra-flat gold substrate, with the same surface coverage ( $S/A$ ) of 2.56, for each independent experimental section using the synergetic approach described in the materials and methods section (see chapter 4) and tested for  $\beta 2m$

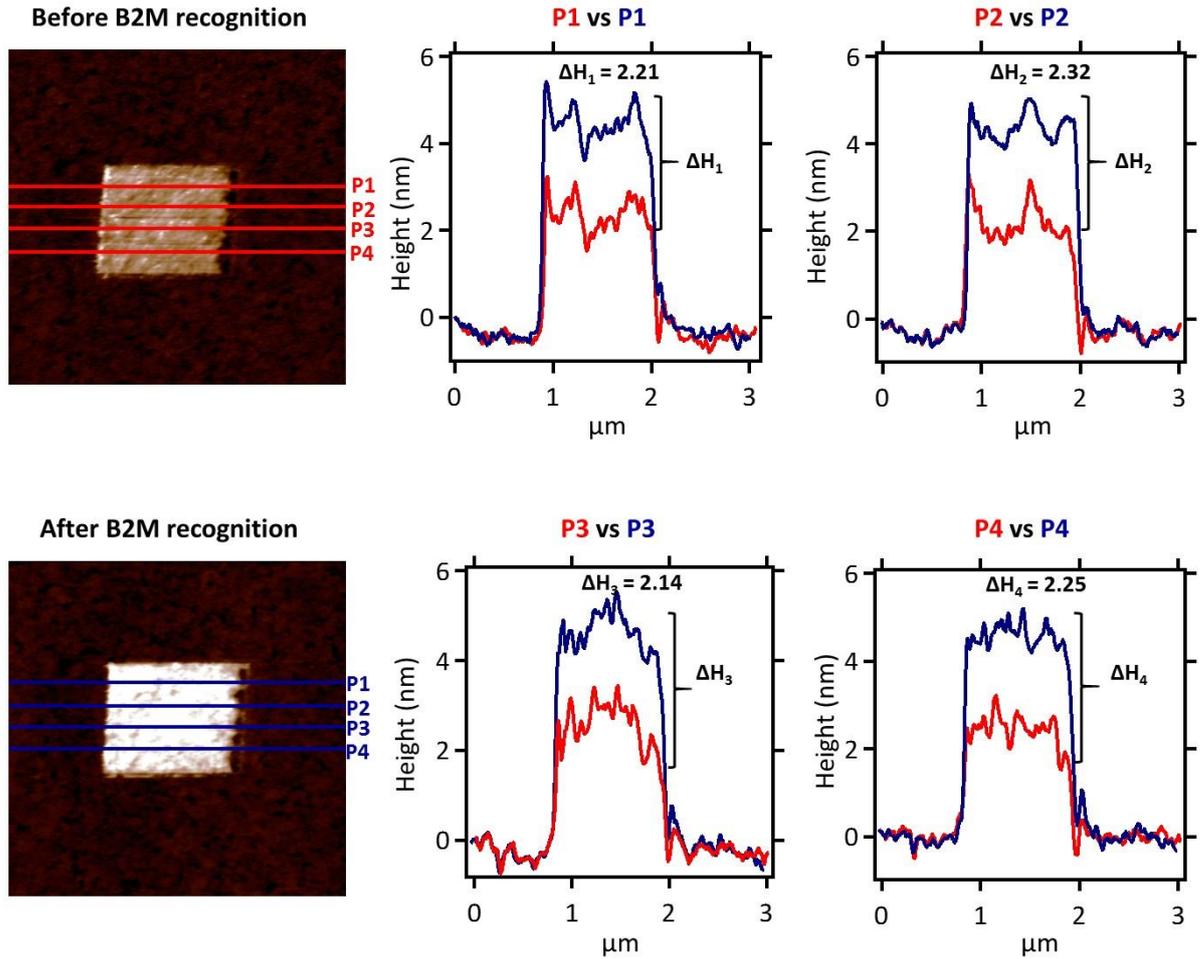
recognition at different concentrations. However, only DNA-pep461 assemblages with surface coverage  $S/A = 2.56$  were considered for the determination of dissociation constant, because this surface coverage yields a consistent and reproducible binding signal (differential height) for all concentration of  $\beta 2m$ .

In Fig. 3.5a the DNA-pep461 assemblage was tested towards the recognition of  $\beta 2m$  proteins in a 50  $\mu M$  solution of  $\beta 2m$ , and the side by side topographic height profile measurement in (Fig. 3.5c) shows no change in the height of the assemblage after  $\beta 2m$  recognition assay. However, in Fig. 3.5d which is another assemblage shows a small increase in height after its interaction with 80  $\mu M$  solution of  $\beta 2m$  as shown in Fig. 3.5f. Lastly, the assemblage in Fig. 3.5g shows a higher incremental height after its exposure to 100  $\mu M$  solution of  $\beta 2m$  as shown in Fig. 3.5i. In addition to the quantitative measurements, we also observed a qualitative result in term of the changes in the brightness and the contrast of assemblages (as show in the assemblages in Fig. 3.5e and 3.5h) after successful  $\beta 2m$  recognition by the DNA-pep461 assemblages. The scan size of the micrograph is 3  $\mu m$  X 3  $\mu m$  while the size of DNA-pep461 assemblage is 1  $\mu m$  X 1  $\mu m$ . This is applicable to all AFM micrographs in Fig. 3.5 and 3.6.



**Figure 3.5:** AFM-based quantitative approach for analysing surface-based DNA-pep461 assemblages before and after  $\beta$ 2m recognition at different target ( $\beta$ 2m) concentration. Three DNA-pep461 nano-assemblages generated in different experimental sections were tested for  $\beta$ 2m recognition at different concentration of proteins, 50  $\mu$ M, 80  $\mu$ M, and 100  $\mu$ M. AFM topographic height measurement was employed in each case, to determine the differential height, which is the binding signal for the detection of  $\beta$ 2m in each case, this is shown in (c), (f) and (i) respectively.

To determine the dissociation constant of DNA-pep461 assemblages towards  $\beta$ 2m protein, the differential height ( $\Delta H$ ), that is the change in height between the height of the DNA-pep461 assemblage before  $\beta$ 2m recognition and the height of the same assemblage after  $\beta$ 2m recognition assay, was measured as shown in Fig. 3.6.



**Figure 3.6:** An example of a subsequent height measurement procedure employed in determining the average height difference across the topographic AFM micrograph of DNA-pep461 assemblage before and after  $\beta$ 2m recognition (100  $\mu$ M concentration). To determine the differential height value for each plot, we employed peak-to-peak subtraction of the height waves for each plot using IGOR-waves analysis. Followed by statistical analysis of the output wave from the peak-to-peak wave subtraction. The same analysis was employed for all the four plots. This yields the average value for  $\Delta H_1$ ,  $\Delta H_2$ ,  $\Delta H_3$ , and  $\Delta H_4$  with their corresponding standard deviation. Lastly, the mean  $\Delta H$  ( $\overline{\Delta H}$ ) which is 2.23 nm in this case, becomes the binding signal in this experiment.

The procedure described in Fig. 3.6 was subsequently done for each DNA-pep461 assemblage tested for  $\beta$ 2m recognition at different  $\beta$ 2m concentration. For each concentration, the average differential height  $\overline{\Delta H}_i$  for the  $i$ -th experiment was calculated as follows;

$$\overline{\Delta H}_i = \frac{\Delta H_1 + \Delta H_2 + \Delta H_3 + \Delta H_4}{4} \quad (\text{iii})$$

Where  $\Delta H_1, \Delta H_2, \Delta H_3$ , and  $\Delta H_4$  are the differential heights for each plot in Fig. 3.6 obtained from peak-to-peak subtraction of the profiles.

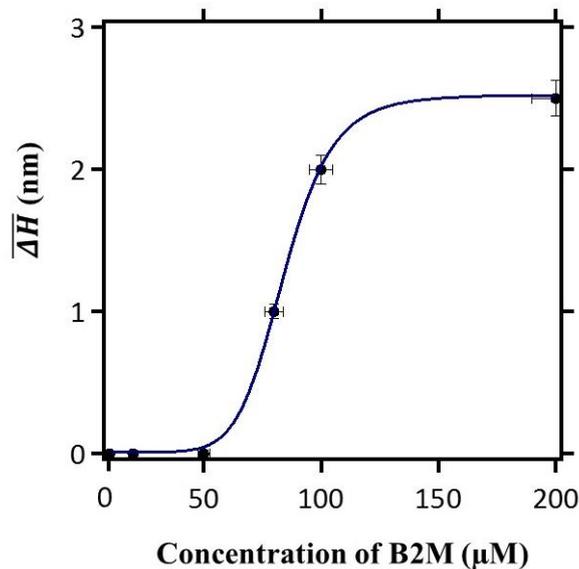
Lastly, the overall average differential height ( $\overline{\Delta H}$ ) obtained for at least three independent experiments for a given  $\beta 2m$  concentration was considered.

Mathematically, we calculated overall average differential height as follows;

$$\overline{\Delta H} = \frac{1}{n} \sum_{i=1}^n \overline{\Delta H}_i \quad (\text{iv})$$

Where  $\overline{\Delta H}_i$  is the average differential height for experiment  $i$  (@ 100  $\mu\text{M}$  of  $\beta 2m$ ), and  $n$  is the number of experiments performed.

These calculations were done for all the DNA-pep461 assemblages tested for  $\beta 2m$  recognition at different concentration and the overall  $\Delta H$ (s) were plotted as a function of  $\beta 2m$  concentration in Fig. 3.7. The experimental data were fitted with Hill's equation using IGOR PRO 6.3A analysis software.



**Figure 3.7:** Graph of overall differential height associated with  $\beta 2m$ -cyclic peptide interactions as a function of  $\beta 2m$  concentration. The experimental data (closed black spheres) was fitted using Hill's equation.

### 3.4 Conclusion

Overall, we have employed a set of computationally evolved peptides showing promising binding affinities towards the solvent exposed binding site on  $\beta 2m$  to immobilize  $\beta 2m$  on a gold surface. In more detail, we have utilized a combination of nanografting,

DNA-directed immobilization and AFM differential height measurements, to characterize in a label-free way the binding affinity of a peptide (pep461) towards  $\beta$ 2m.

We have first shown that the surface coverage of DNA-arrays before their transformation into DNA-peptide arrays through hybridization can influence the succeeding biochemical reaction, if the surface density is above a certain threshold. Following this observation, we generated the DNA arrays with constant optimum surface density ( $S/A = 2.56$ ), which then allows the successful functions of the arrays for hybridization process and protein recognition.

By fitting the experimental results with Hill's equation, as represented on the plot in Fig. 3.7, which is the overall differential relative height as calculated using equation (iv) with respect to the concentration of  $\beta$ 2m, we obtained the binding affinity of DNA-pep461 toward  $\beta$ 2m to be  $84.1 \pm 0.4 \mu\text{M}$ . Concurrently, the binding affinity obtained via ITC technique (see Appendix B) is  $62 \mu\text{M}$ . However, the binding affinity obtained in SPR experiments is of 5-times ( $\sim 400 \mu\text{M}$ ) lower than that of AFM and ITC. The difference in AFM and SPR results could be because of change in the choice of probe; in AFM experiments, peptide was the probe on surface while in SPR experiments  $\beta$ 2m was the probe on the surface without taking care that the protein orientation be constant and that clearly can influence the results.

In our experiments, however, we cannot quantitatively measure the surface density ( $\text{mol} / \mu\text{m}^2$ ) of the DNA-peptide nanoassemblage, we could only use the calibration scale of  $S/A$  to tune the surface density of DNA arrays during the nanografting process. The possible technique that can circumvent this challenge is Quartz Crystal Microbalance (QCM), and this can be integrated into the AFM experimental setup for the future experiments.

### 3.5 References

1. Hong Enriquez, R. P.; Pavan, S.; Benedetti, F.; Tossi, A.; Savoini, A.; Berti, F.; Laio, A. *Journal of Chemical Theory and Computation* **2012**, 8, (3), 1121-1128.
2. Gladich, I.; Rodriguez, A.; Hong Enriquez, R. P.; Guida, F.; Berti, F.; Laio, A. *The Journal of Physical Chemistry B* **2015**, 119, (41), 12963-12969.
3. Russo, A.; Scognamiglio, P. L.; Enriquez, R. P. H.; Santambrogio, C.; Grandori, R.; Marasco, D.; Giordano, A.; Scoles, G.; Fortuna, S. *PloS one* **2015**, 10, (8), e0133571.
4. Verdone, G.; Corazza, A.; Viglino, P.; Pettirossi, F.; Giorgetti, S.; Mangione, P.; Andreola, A.; Stoppini, M.; Bellotti, V.; Esposito, G. *Protein science* **2002**, 11, (3), 487-499.
5. Saper, M.; Bjorkman, P.; Wiley, D. *Journal of molecular biology* **1991**, 219, (2), 277-319.
6. Fung, E. T. *Clinical chemistry* **2010**, 56, (2), 327-329.
7. Locatelli, F.; Mastrangelo, F.; Redaelli, B.; Ronco, C.; Marcelli, D.; La Greca, G.; Orlandini, G. *Kidney international* **1996**, 50, (4), 1293-1302.
8. Becker, J. W. & Reeke, G. N. Three-dimensional structure of beta 2-microglobulin. *Proc. Natl. Acad. Sci. U. S. A.* **82**, 4225-9 (1985).
9. Collis, A. V.; Brouwer, A. P.; Martin, A. C. *Journal of molecular biology* **2003**, 325, (2), 337-354.
10. Russo, A.; Aiello, C.; Grieco, P.; Marasco, D. *Current medicinal chemistry* **2016**, 23, (8), 748-62.
11. Sugita, Y.; Okamoto, Y. *Chemical physics letters* **1999**, 314, (1), 141-151.
12. Daura, X.; Gademann, K.; Jaun, B.; Seebach, D.; van Gunsteren, W. F.; Mark, A. E. *Angewandte Chemie International Edition* **1999**, 38, (1-2), 236-240.
13. Iba, Y. *International Journal of Modern Physics C* **2001**, 12, (05), 623-656.
14. Hess, B.; Kutzner, C.; Van Der Spoel, D.; Lindahl, E. *Journal of Chemical Theory and Computation* **2008**, 4, (3), 435-447.
15. Schafmeister, C.; Ross, W.; Romanovski, V. *University of California San Francisco: San Francisco, USA* **1995**.
16. Trott, O.; Olson, A. J. *Journal of computational chemistry* **2010**, 31, (2), 455-461.
17. Goddard, T.; Kneller, D., San Francisco: University of California. In SPARKY.

18. Ganau, M. *et al.* A DNA-based nano-immunoassay for the label-free detection of glial fibrillary acidic protein in multicell lysates. *Nanomedicine Nanotechnology, Biol. Med.* **11**, 293–300 (2015).
19. Bano, F.; Fruk, L.; Sanavio, B.; Glettenberg, M.; Casalis, L.; Niemeyer, C. M.; Scoles, G. *Nano letters* **2009**, 9, (7), 2614-2618.
20. Bosco, A., Ambrosetti, E., Mavri, J., Capaldo, P. & Casalis, L. Miniaturized Aptamer-Based Assays for Protein Detection. *Chemosensors* **4**, 18 (2016).
21. Solulink. Antibody-Oligonucleotide All-In-One Conjugation Kit Technical Manual. *Solulink Man. A-9202-001*, 1–38 (2011).
22. Liu, M. & Liu, G. Y. Hybridization with nanostructures of single-stranded DNA. *Langmuir* **21**, 1972–1978 (2005).
23. Mirmomtaz, E. *et al.* Quantitative study of the effect of coverage on the hybridization efficiency of surface-bound DNA nanostructures. *Nano Lett.* **8**, 4134–4139 (2008).
24. Bosco, A. *et al.* Hybridization in nanostructured DNA monolayers probed by AFM: theory versus experiment. *Nanoscale* **4**, 1734 (2012).

## 4.0 Results and discussion on surface-bound high affinity binders for protein recognition\*

*In this chapter, all the experimental validation based on atomic force microscopy, to determine the sensitivity of two peptides to be linked in the future, which bind to two distinct binding sites on  $\beta 2m$ , were performed by me.*

Here we present preliminary results aiming at the design of surface-bound bidentate artificial receptors. The receptors are inspired by the concept of “chelation” in which two binders, designed each to identify a distinct binding site on the protein surface collectively achieve higher sensitivity. A new design protocol has been put forward for the peptide design. The protocol relies on the identification of possible binding sites of a set of  $\beta 2m$  conformations obtained by the webtool Peptimap<sup>1</sup>. A new set of peptides were generated and screened. We then combined nanografting<sup>2</sup> and DNA-directed immobilization<sup>3</sup> as experimental validation approach to immobilize the cyclic peptide (tagged pep331) on an ultra-flat gold substrate, and through AFM height measurements, we quantitatively characterize the binding and sensitivity of DNA-pep331 assemblage towards  $\beta 2m$ . The same approach has been further utilized to characterize the sensitivity of a formerly designed cyclic peptides<sup>4</sup> (here labeled pep381). These two peptides, pep381 and pep331, are then shown successful in binding to two distinct solvent exposed sites on  $\beta 2m$ .

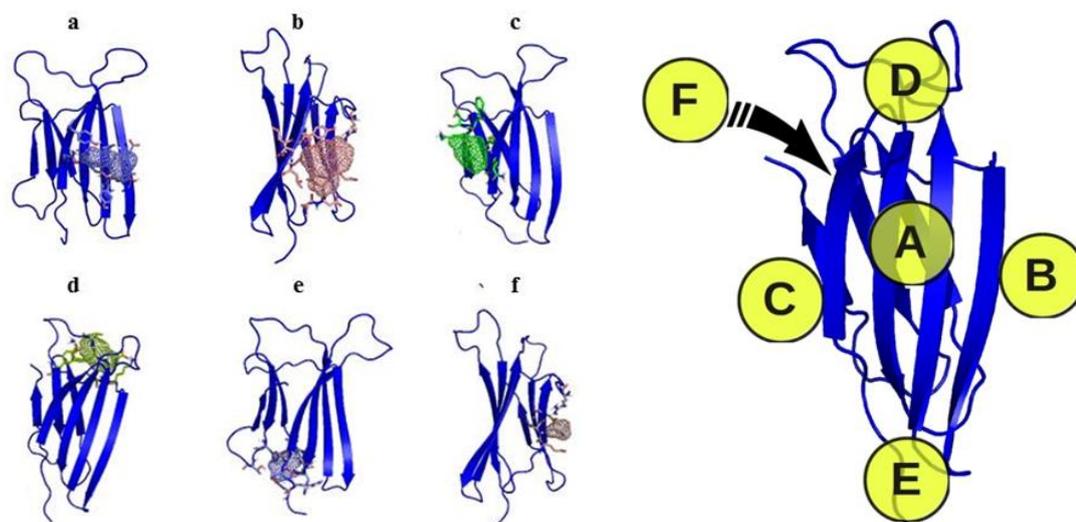
### 4.1 Binding site identification, peptide design, and screening

We run 20ns of molecular dynamics simulation of the free protein. From the trajectory, we selected 12 representative snapshots. These configurations were analyzed by Peptimap to identify possible druggable site on the protein surface. Fig. 4.1 shows the six identified binding sites, while the ranking for each protein snapshot is shown in Table 4.0. According to these results, the binding site located in the region B is the most probable,

---

\* Manuscript in preparation

since it appears in the 1st position in every snapshot. Otherwise, the region A (the one chosen in the work of Ref.4 appears as 2nd and 3rd most probable binding site in most of the snapshots. The third most probable region for allocating a peptide binding site is the region C, which occurs most likely as a 4<sup>th</sup> possible binding site, while occurring also as 2nd and 3<sup>rd</sup> possible site for some snapshots.

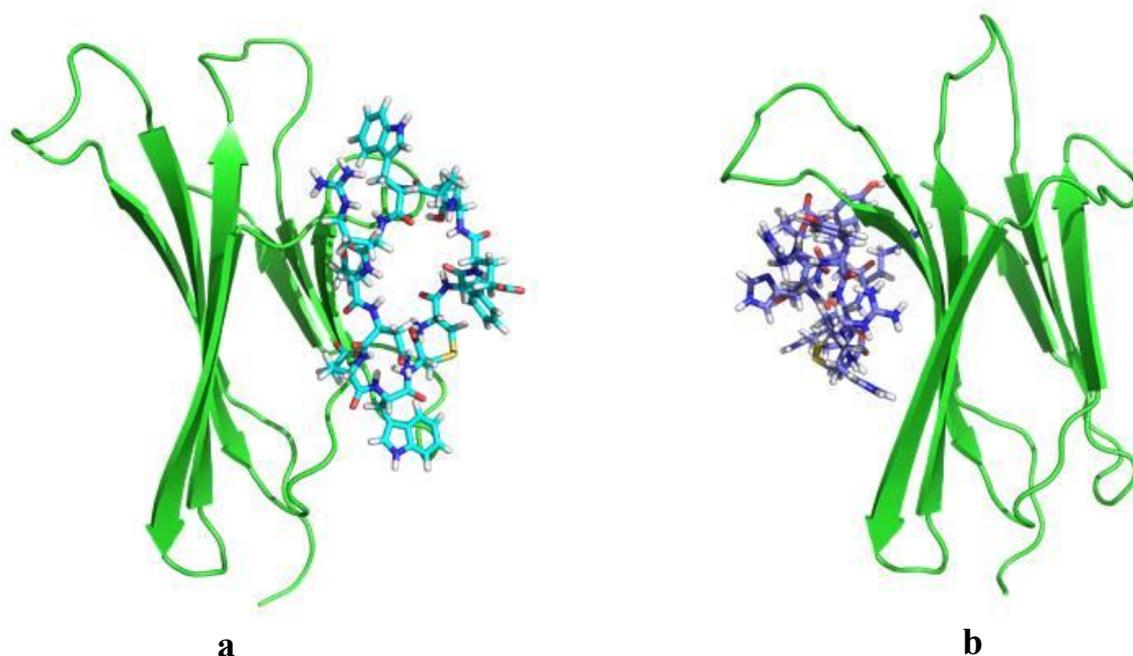


**Figure 4.1:** Peptimap graphical output representing a selection of the identified binding sites. These are all located in the regions (a) A, (b) B, (c) C, (d) D, (e) E and (f) F of  $\beta$ 2m and the representation on the far right shows the schematic of the different labeled regions of  $\beta$ 2m.

**Table 4.0:** Ranking of different binding sites obtained for each  $\beta$ 2m conformation.

B2M snapshots	Binding Site Ranking					
	1 <sup>st</sup>	2 <sup>nd</sup>	3 <sup>rd</sup>	4 <sup>th</sup>	5 <sup>th</sup>	6 <sup>th</sup>
0.0 ns	B	B	B	C	B	A
1.2 ns	B	C	A	A	C	D
2.5 ns	B	A	D	C	A	F
3.7 ns	B	C	E	E	C	A
5.0 ns	B	D	A	C	C	A
7.5 ns	B	C	A	A	C	B
10.0 ns	B	A	E	F	D	C
11.2 ns	B	A	A	C	C	E
12.5 ns	B	A	C	F	A	E
15.0 ns	B	B	C	D	A	D
17.5 ns	B	A	E	C	A	A
20.0 ns	B	C	A	C	A	A

We chose the binding site in the region B as novel binding site for the design of a second set of peptides (see Fig. 4.2a), while we selected from our previous work a peptide for the binding site in the region A (Fig. 4.2b). Once again, the algorithm has been run twice, leading to two sets of peptides. The generated peptides have been then screened by means of molecular dynamics simulations to identify those with the lowest binding energy and the closest distance from their binding site, which has been identified as optimum indicators for the peptide screening in the work of Chapter 3. One peptide, which will be referred to as pep331 in the rest of the chapter, has been then chosen for further characterization (Fig. 4.2a). Hence, the Fig. 4.2 represents the  $\beta$ 2m-peptide complexes for the two distinct binding sites.



**Figure 4.2:** Simulation snapshot of (a) the pep331 and (b) pep381 with  $\beta$ 2m, each binding a different solvent exposed binding site on  $\beta$ 2m.

#### 4.2 Experimental validation of $\beta$ 2m -peptide binding via atomic force microscopy

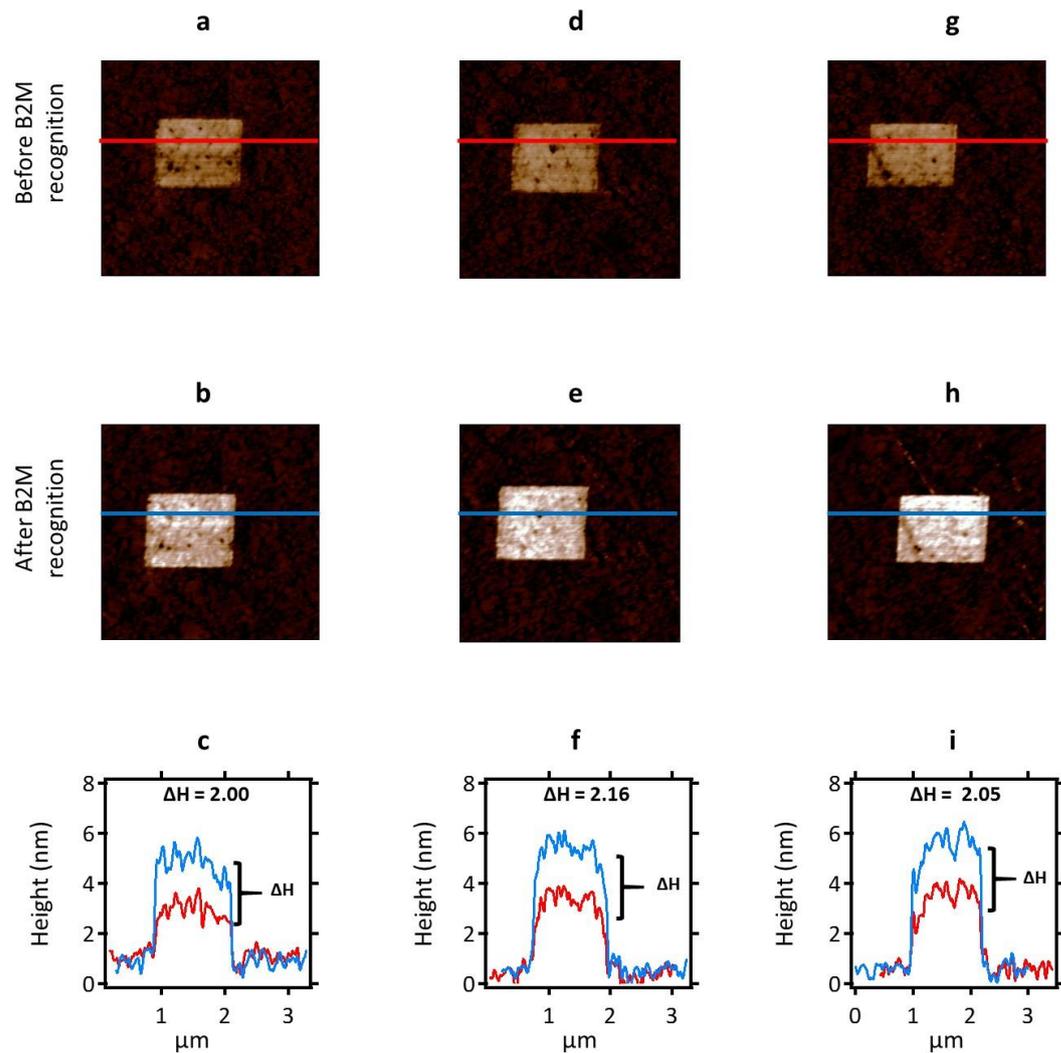
Prior to the immobilization steps, the two cyclic peptides, pep381 and pep331, were conjugated with two distinct DNA sequences (see the materials section for details) using the NHS-ester-primary amine conjugation chemistry<sup>4,9</sup>. Subsequently, we utilized DNA-directed immobilization (DDI)<sup>3,10</sup> combined to Nanografting<sup>2,4</sup> to form nanopatches of spatially-oriented DNA-cyclic peptides able to recognize the solvent-exposed binding site on  $\beta$ 2m.

As described in chapter 3, we firstly grafted thiol-derivatized ssDNA (i.e. CD1 or CD2 with respect to the peptide we intended to use), with area  $1 \times 1 \mu\text{m}^2$  and constant surface coverage<sup>4,11</sup> ( $S/A = 2.56$  and concentration of  $1 \mu\text{M}$ ) within an ethylene glycol terminated monolayer passivated on flat gold substrate. Following the successful grafting procedure (see the details in material and method section) as indicated by the fast imaging at low loading force ( $\sim 0.2 \text{ nN}$ ) of the  $20 \times 20 \mu\text{m}^2$  scan size, which contains the grafting section, we then carried out hybridization step, by introducing  $200 \text{ nM}$  solution of the DNA-peptide conjugate (e.g. D1-pep381 or D2-pep-331 with respect to the grafted complementary), this step lasted for one hour at  $37 \text{ }^\circ\text{C}$ . At the end of the hybridization step, we imaged the grafted section at low loading force and measured the relative initial height before carrying out the binding assays with  $\beta 2\text{m}$ . The binding assay lasted for 1 hour at  $25 \text{ }^\circ\text{C}$ .

Lastly, we performed the topographic imaging again after the binding assay, and measured the relative final height with respect to the surrounding bio-repellent monolayer of alkyl thiol. The difference in relative height with respect to the before  $\beta 2\text{m}$  and after  $\beta 2\text{m}$  recognition becomes the signature of successful recognition of  $\beta 2\text{m}$  by the patterned DNA-peptide nanoassemblage.

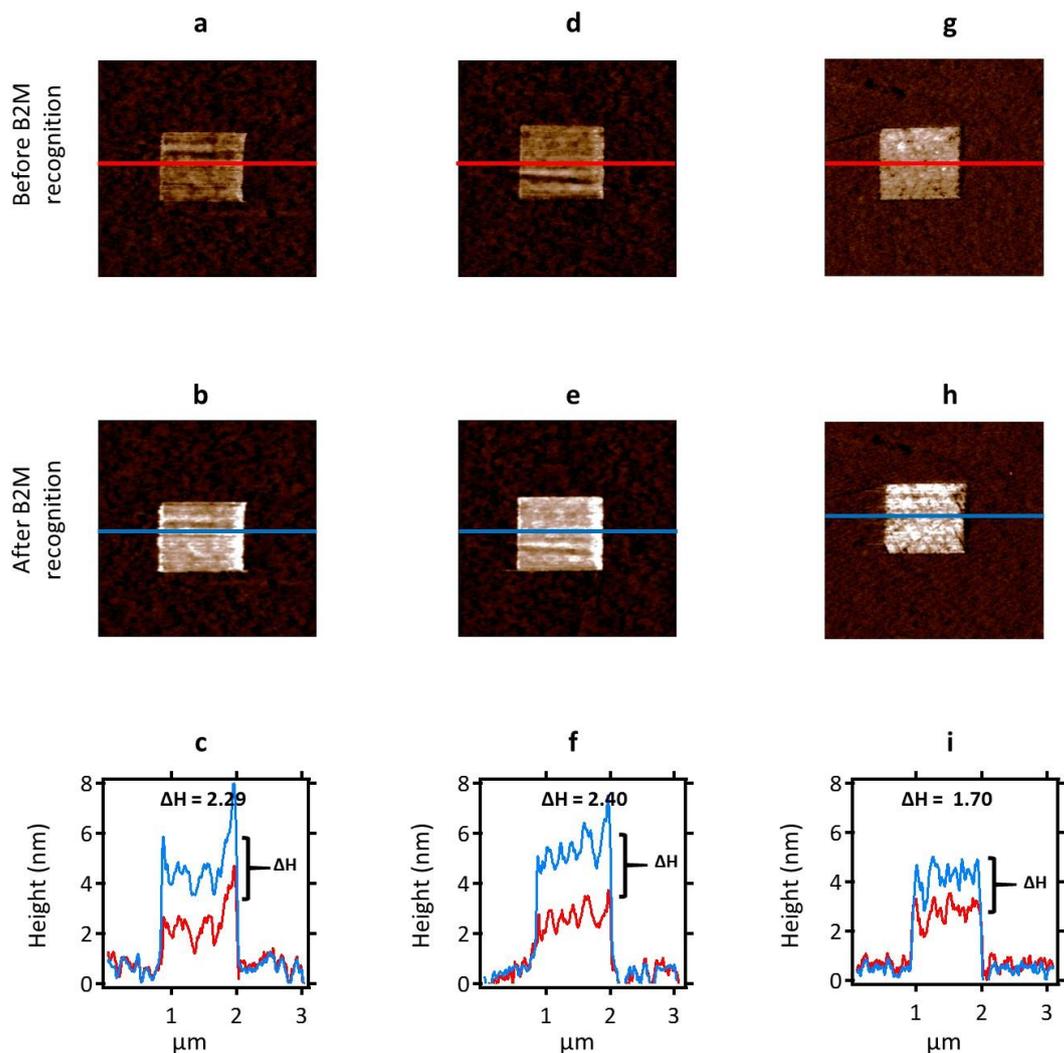
Starting with cyclic pep331, we employed the aforementioned approaches to construct spatially-oriented pep331 terminated DNA nanoassemblage (DNA-pep331). Each column in Fig. 4.3 represents the three-step quantitative analysis based on AFM; 1) AFM topographic height micrograph with line profile (red colour) before  $\beta 2\text{m}$  recognition, followed by 2) AFM topographic micrograph of the same after  $\beta 2\text{m}$  recognition, and lastly 3) the relative height profiles, indicating the recognition of  $\beta 2\text{m}$  by DNA-pep331 assemblages, as evident in the change in the relative height profiles. In Fig. 4.3 a, d, g we have three DNA-pep331 nanoassemblages generated independently with the same surface coverage ( $S/A = 2.56$ ), each was tested for  $\beta 2\text{m}$  recognition in the standard solution with  $38 \mu\text{M}$  concentration of  $\beta 2\text{m}$ . Qualitatively, the colour maps of these three DNA-pep331 nanoassemblages a, d, and g change after the recognition of  $\beta 2\text{m}$ , see Fig. 4.3 b, e, g, indicating increase in relative height after the  $\beta 2\text{m}$  binding height; brighter stands for high height and darker stands for low height.

Quantitatively, the AFM topographic height profiles in 4.3 c, f, and i, shows that the relative initial height of the three DNA-pep331 nanoassemblage in a, d, g increases after the recognition of  $\beta 2\text{m}$ .

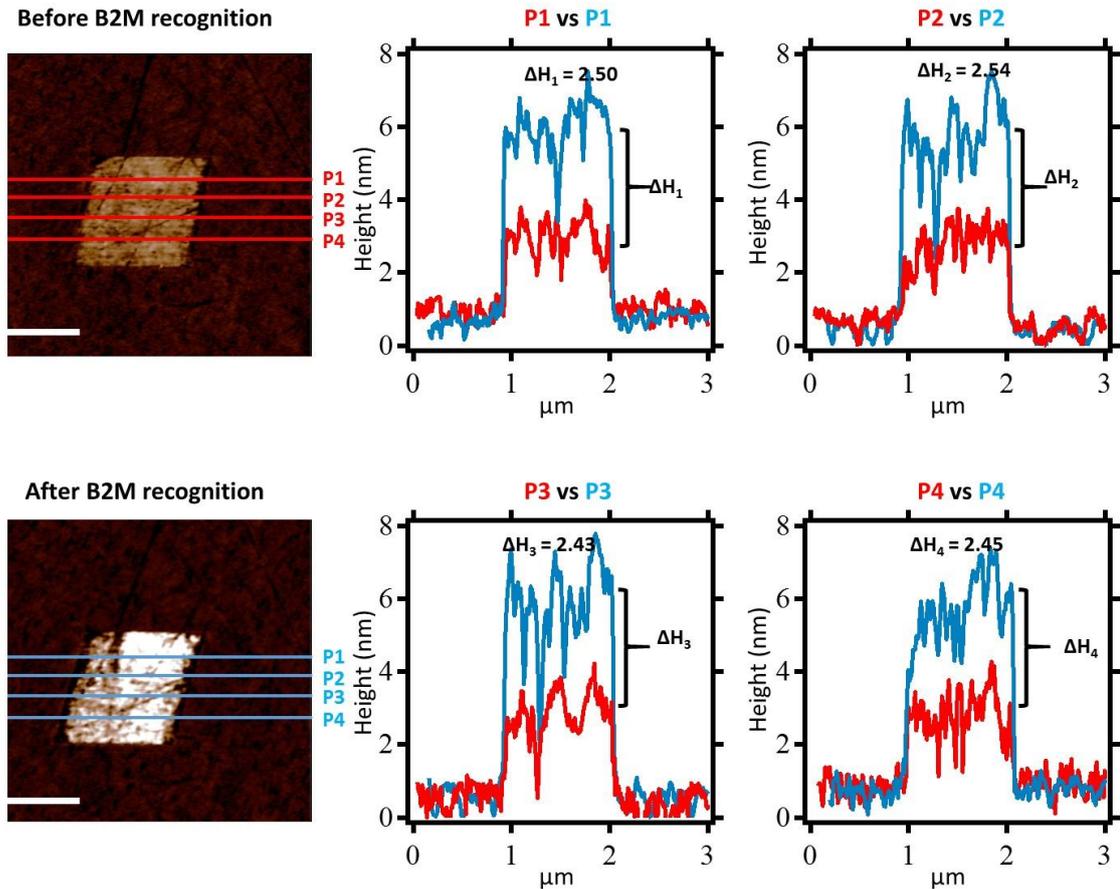


**Figure 4.3:** Using AFM topographic height measurement to detect the recognition of DNA-peptide assemblages (DNA-pep331) towards  $\beta$ 2m at 38  $\mu$ M concentrations. Each column is divided into three sections; Before  $\beta$ 2m recognition, after  $\beta$ 2m recognition and the relative height of the peptide-terminated assemblage with respect to before and after  $\beta$ 2m recognition. The scan size of all the AFM micrographs is 3 x 3  $\mu$ m<sup>2</sup>.

Additionally, we also constructed cyclic peptide381 terminated DNA nanoassemblages (DNA-pep381), using the same approach for generating those of DNA-pep331 and we then tested the sensitivity of the DNA-pep381 nanoassemblage towards the second binding site on  $\beta$ 2m at 7  $\mu$ M concentration of  $\beta$ 2m solution (see Fig. 4.4).

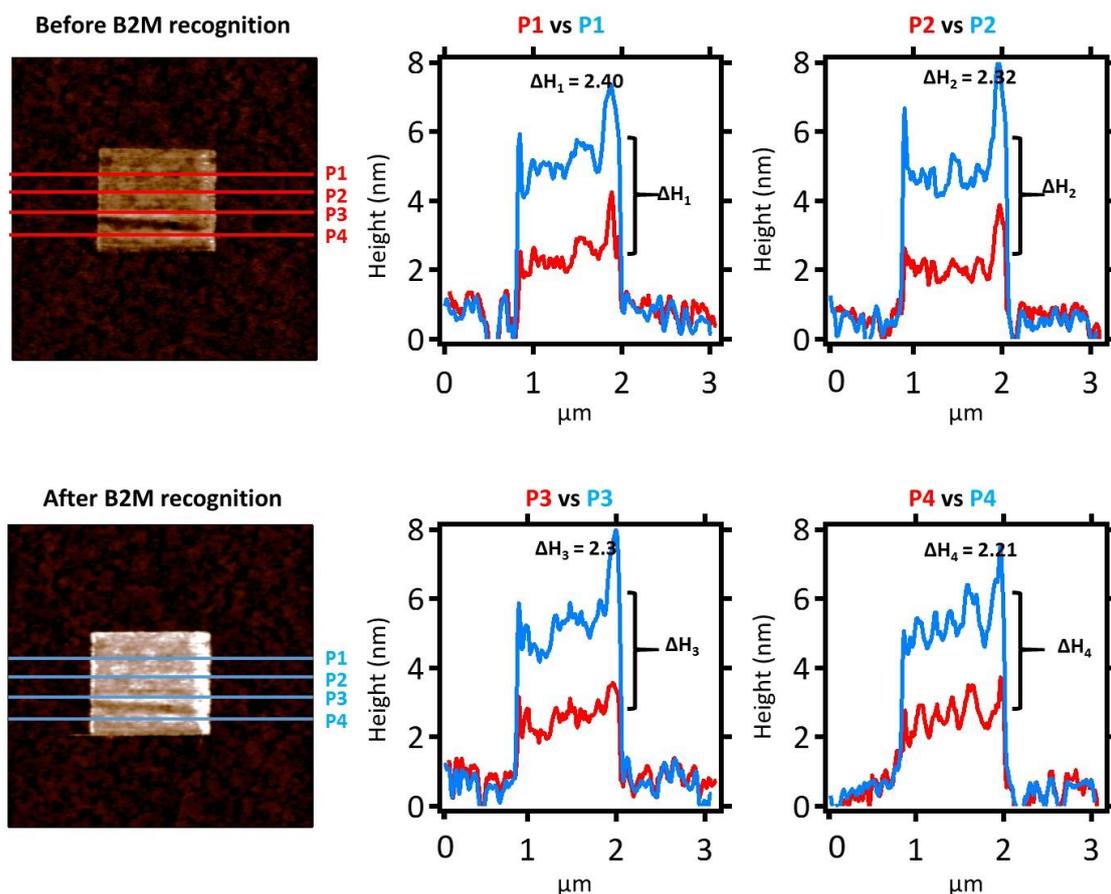


**Figure 4.4:** Recognition of  $\beta$ 2m by DNA-pep381 assemblages at the 7  $\mu$ M concentration of target solution. The recognition was detected via label-free AFM topographic height measurements. The first row shows the AFM topographic height micrographs of three DNA-pep381 nanoassemblages a, d, and g before  $\beta$ 2m binding assay. Followed by the second row, which are the same AFM micrographs in a, d, g but after the recognition of  $\beta$ 2m at 7  $\mu$ M concentration of  $\beta$ 2m solution. Lastly, the third row shows the plots of AFM height profiles of the three DNA-pep381 nanoassemblage, corresponding to the relative initial height of DNA-pep381 before (red colour) and after (light blue colour)  $\beta$ 2m recognition. The changes in relative heights ( $\Delta$ Hs) for the three DNA-pep381 nanoassemblages were calculated using the peak-to-peak wave subtraction<sup>4</sup> in IGOR PRO work space.



**Figure 4.5:** Shows the serial height measurement procedure with one-to-one profile mapping employed in determining the average height difference across the topographic AFM micrographs of DNA-pep331 assemblage before and after  $\beta 2m$  recognition ( $7 \mu M$  concentrations). Scale bar, 750 nm.

Furthermore, in Fig. 4.5 and 4.6 we clearly show the procedures of AFM height measurements, through which we investigate the binding or recognition over the whole peptide-DNA assemblages. Firstly, we measured 4-5 height profiles over the AFM micrograph of the peptide-DNA nanoassemblage with respect to the before and after  $\beta 2m$  recognition at any given concentration of the  $\beta 2m$ , (see Fig. 4.5 and 4.6). Next, using the Igor objects; waves, tables and wavestats function, we collate each relative height wave tallying with before  $\beta 2m$  and after  $\beta 2m$  into distinct column, followed by peak-to-peak subtraction of height waves, then wavestats of the resultant wave, which results into statistics of the wave. These procedures are performed for all the profiles (four profiles) in Fig. 4.5 and 4.6 respectively.



**Figure 4.6:** Subsequent height measurement procedure employed in determining the average height difference across the topographic AFM micrographs of DNA-pep381 assemblage before and after  $\beta$ 2m recognition (at 7  $\mu$ M concentration). The scan size of the AFM micrographs is 3 x 3  $\mu$ m<sup>2</sup>.

From the statistics data, we obtained the differential height changes with respect to each profile; for example,  $\Delta H_1 = 2.50$  nm for DNA-pep331 nanoassemblage with respect to line profile 1 (**P1** vs. **P1**) on the two AFM micrographs.

Lastly, we calculated the mean of all the differential heights ( $\Delta H_{i = 1, 2, 3, 4}$ ) and this become the average differential height that indicate the binding and recognition of  $\beta$ 2m. From Fig. 4.5 and 4.6, the average differential heights are 2.48 nm and 2.31 nm respectively, indicating the recognition of  $\beta$ 2m by DNA-pep331 and DNA-pep381 nanoassemblage at 7  $\mu$ M concentration of  $\beta$ 2m solution. Also, we noticed a high surface roughness across the DNA-pep331 nanoassemblage (see Fig. 4.5), which remain unchanged before and after the  $\beta$ 2m recognition. This surface roughness only occurs in this experimental section, other experimental section for DNA-pep331 nanoassemblage (see Fig. 4.3) is without such roughness. Therefore, this observation could be as a result of the surface roughness of the

gold substrate, which in turn reflected across the DNA-pep331 nanoassemblages after it was generated. Likewise, this can be associated with the differences in the average differential heights as calculated.

Unequivocally, from the results observed in Fig. 4.5 and 4.6, the two cyclic peptides, pep331 and pep381 recognise their respective binding sites on  $\beta$ 2m with high sensitivity. Also, the results further confirmed that the bioactivities of cyclic peptides are un-hindered by either nanografting process or bio-conjugation procedures.

### 4.3 Discussion and conclusion

In this chapter, our collaborators have introduced a new computer algorithm that generates several putative binding sites for the protein of interest, in our case Beta-2-microglobulin. Followed by screening and ranking of the peptide residues in relation to (1) their binding energies, (2) distances between the centre of mass of the  $\beta$ 2m and the peptide, (3) root mean deviations of the binding site(BS) with respect to the peptides in different conformations and (4) the hydrophobic surface area of the peptides, using the combination of Vina, LIGPLOT and GROMACS (see Appendix).

Furthermore, we synergized DNA-directed immobilization and AFM-based nanolithography technique termed nanografting, to construct peptide-functionalized DNA nanoassemblages, with spatially-oriented conformation, which binds to the solvent exposed binding sites on  $\beta$ 2m. Additionally, we employed AFM height measurement to quantitatively characterize the recognition of  $\beta$ 2m by peptide-DNA nanoassemblages, using an automated analysis procedure (without user interference) based on Igor procedure. With such strategies, we show the high sensitivity of two distinct cyclic peptides, pep381 and pep331, each with different 10 amino-acids sequences, binding to two different binding sites on  $\beta$ 2m. According to this results here, both peptides bind  $\beta$ 2m with maximum sensitivity of  $7\mu\text{M}$ . However, we need to carry out more measurements in the lower concentration regime, to pin-down precisely the high limit of dissociation constant. Comparing our results with those in chapter 3, we have developed two *in silico* based cyclic peptides with 8-fold higher in sensitivity ( $\sim 7\mu\text{M}$ ) with respect to pep461 that in chapter 3 shown to have a binding constant of  $84\mu\text{M}$ . Additionally our AFM-based validation approach promises to yields more reliable and reproducibility binding sensitivity compared to surface plasmon resonance (SPR, see values in chapter 3), this could be due to the change in configuration, as peptides were immobilized through DDI on surface in the former approach while  $\beta$ 2m

was immobilized on the surface in the SPR experiments. Also, the low sensitivity in SPR could be as a result of high surface density of  $\beta 2m$  on the surface, with random orientation and these could lead to in-accessibility of the targets.

#### **4.4 Materials and methods for the experimental validations for pep461, pep381, and pep331**

Three thiol-derivatized (with C6 linker) single stranded (ss) DNA and their corresponding complementary DNA (with amino-C6-link) were purchased as HPLC purified grade from Biomers GmbH (Ulm, Germany). Top-oligo-ethylene glycol (TOEG 6 (1-mercaptoundec-11-yl) hexa(ethyleneglycol), (HS-((CH<sub>2</sub>)<sub>11</sub>)-(O-CH<sub>2</sub>-CH<sub>2</sub>)<sub>6</sub> -OH)) was purchased from Prochimia Surfaces (Poland), sodium chloride (NaCl), Tris-EDTA, sodium phosphate dibasic (NaH<sub>2</sub>PO<sub>4</sub>), sodium phosphate monobasic (NaHPO<sub>4</sub>), and absolute ethanol (99.8%) were all purchased from Sigma Aldrich. All buffer solution was prepared in ultra-pure water (miliQ-H<sub>2</sub>O), of resistivity 18.2 M $\Omega$ -cm at 25 °C and filtered before use with sterile syringe filter (of 0.22 $\mu$ m pore size). Also  $\beta 2m$  was purchased by SIGMA, while cyclic peptides were purchased as 95% purity from ProteoGenix SAS (Schiltigheim, France). Also, Beta-2-Microglobulin (Human species) was purchased from Sigma Aldrich (Germany)

#### **Instrumentation**

All AFM experiments (starting from scanning, nanografting, and height detection of  $\beta 2m$ -cyclic peptide binding interactions) were carried out with MFP-3D-BIO AFM, (Asylum research, Santa Barbara, CA). For nanografting, we used commercially available pyramidal silicon etched probes NSC 18 /no Al of spring constant of 2.3 N/m (Mikro-Masch, Germany) was used. For contact mode imaging, soft probe (CSC 38 /no Al, spring constant 0.03N/m, Mikro-Masch, Germany) was used.

#### **DNA and peptide sequences**

cD0 (22 bps long): SH-(CH<sub>2</sub>)<sub>6</sub>-5'-CTTCACGATTGCCACTTTCCAC -3'  
D0 (22 bps long) aminolink-C6-5'-GTGGAAAGTGGCAATCGTGAAG-3'  
cD1 (22 bps long): SH (CH<sub>2</sub>)<sub>6</sub>-5'-TTCGGCTCATACTCTGACTGTA-3'  
D1 (22 bps long): amino link-C6-5'-TACAGTCAGAGTATGAGCCGAA-3'  
cD2 (22 bps long): SH (CH<sub>2</sub>)<sub>6</sub>-5'-CTTATCGCTTTATGACCGGACC-3'  
D2 (22 bps long): aminolink-C6- 5'-GGTCCGGTCATAAAGCGATAAG-3'

Cyclic peptide-461(10-amino acids long): [C]YFTKTDRRFQ[C] (with single disulfide bridge at Cys 1-Cys 12)

Cyclic peptide-381(10- amino acids long): [C]RRYSHQHRYR[C] (with single disulfide bridge at Cys 1-Cys 12)

Cyclic peptide331(10- amino acids long): [C]FETAWRQNEW[C] (with single disulfide bridge at Cys 1-Cys 12)

### **DNA-peptide conjugates preparation**

SoluLink's superior bioconjugation method was exploited to prepare peptide-oligonucleotide conjugates in three steps. (i) Peptide modification: a volume that represents 10 mole equivalents of HyNic/mole protein was added to the peptide (concentrated 2.0 mg/mL) and mixed. The reaction was carried out at room temperature for 1.5 hours. (ii) Oligonucleotide modification: the oligonucleotide D0-amine (Biomers GmbH Ulm, Germany) was desalted into nuclease free water using a 5K MWCO VivaSpin diafiltration apparatus and OD/uL concentration at 260 nm was determined (0.4 OD/ $\mu$ L). A volume containing 20 equivalents S-4FB was added to the oligonucleotide solution and incubated at room temperature for 2 hours. Next, the 4FB-modified oligonucleotide was desalted into conjugation buffer (100 mM phosphate, 150 mM NaCl, @ pH 6.0). (iii) DNA-peptide conjugation: taking into account the concentration and the mass of the HyNic-modified peptide to be functionalized and concentration and MSR of 4FB-modified oligonucleotide, volumes of the two components to be mixed are determined; 1/10 volume 10X TurboLink Catalyst Buffer (100 mM aniline, 100mM phosphate, 150 mM NaCl, pH 6.0) was added to the conjugation solution and the reaction was carried out overnight at 4°C. The concentration of DNA-peptide conjugates was determined spectrophotometrically by the absorbance at A354 due to the formation of the chromophoric conjugate bond. The aforementioned conjugation method was used for the three cyclic peptides, resulting into pep461-D0, pep381- D1 and pep331-D2 conjugates respectively.

### **Preparation of ultra-flat gold substrate**

Ultra-flat gold substrates were prepared as described in the past publications from our group<sup>11,12</sup>. Briefly, a sequential deposition of gold was employed using electron beam evaporation. Firstly, gold was deposited at the rate 0.05nms<sup>-1</sup> until a film of 5nm was obtained, then the rate of evaporation was increased to 0.1nms<sup>-1</sup> until a film of 100nm thick was formed on freshly cleaved mica (Mica New York Corp., clear ruby muscovite) at 10<sup>-</sup>

<sup>6</sup>mbar, at room temperature but in ultra-high vacuum environment. The planar gold sheet of (111) crystallographic plane obtained on mica was sliced into few millimetre squares (approximately 5x5 mm<sup>2</sup>) in area using Stanley-199 blade. To transfer the ultra-flat gold surface from mica to the polished side of smaller squares (smaller than sliced gold on a mica sheet) p-doped silicon wafers, a drop of SU-8 photoresist adhesive (negative tone photoresist, MicroChem, USA) was evenly dispensed on the polished side of silicon. Then a sandwiched was obtained, silicon-gold-mica, by impressing the polished section of silicon on the gold section of mica. All silicon-gold-mica sandwich squares were cured at 130°C for at least 24 hours. The samples were allowed to cool down to room temperature without any external cooling system and this was done to avoid thermal stresses that can result into gold film detachment from the mica substrate. Without any further surface treatment, the samples were stored at room temperature, ready to be used for self-assembled monolayer preparation.

#### **Preparation of top-oligo-ethylene-glycol SAM (TOEGSAM) on the ultra-flat gold substrate**

An ultra-flat gold surface was obtained from silicon-gold-mica sandwich by mechanical stripping the mica substrate from silicon-gold-mica sandwich. This was immersed and incubated in the solution of 100µM of top-oligo-ethylene glycol (TOEG, 6, (HS-((CH<sub>2</sub>)<sub>11</sub>) - (O-CH<sub>2</sub>-CH<sub>2</sub>)<sub>6</sub> -OH), Prochimia and Sigma Aldrich) in absolute ethanol and 1M NaCl, TE1X (10-mM Tris-HCl, 1-mM EDTA, pH 7.2 at 25°C in Milli-Q water) for 6-hours. The TOEG template then serves as bio-repellent for aspecific adsorption of macromolecules on surface. After the incubation time, serial rinsing was performed in both ethanol and 1M NaCl, Te1X buffer to remove any physically-adsorbed contaminants.

#### **Nanografting of thiol-modified ssDNA in contact mode**

Firstly, TOEGSAM passivated gold substrate was adhered firmly on a home-made liquid cell. The ssDNA solution (1µM ssDNA, 1M NaCl, TE1X) was evenly dispensed on the SAM and it was transferred onto the AFM X-Y scanner stage. Secondly, we performed preliminary scanning in tapping mode at low force to obtain the topographic image of the TOEGSAM surface. This allow us to have a survey of the surface and select a flat and clean section for nanografting. At high force ( $\approx 120$  nN),  $1 \times 1\mu\text{m}^2$  was selected for grafting process; the prior alkanethiols were exchanged for thiol-modified ssDNA molecules in

solution. After grafting process, a  $20 \times 20 \mu\text{m}^2$  section that contains the grafted patches was scanned at very low force (high set point) in tapping mode.

### **Hybridization with complementary DNA-peptide conjugates**

After the nanografting process, the sample was thoroughly rinsed with 1M NaCl, TE1X buffer (DNA-free buffer), followed by hybridization procedure with 200 nM of oligo-peptide conjugates solution in 1 M NaCl, TE1X @ pH 7.2 and the hybridization process took place at 37°C for 1-hour. Next, the sample was copiously rinsing with 1M NaCl, TE1X buffer (DNA-free buffer). This is to remove any un-bound molecules and other excess molecules that were not part of the hybridization. Note 100 nM concentration of DNA-pep461 was used for hybridization in chapter 3. At the end of hybridization process before  $\beta$ 2m binding assays, we performed AFM topographic imaging at low force ( $\sim 0.2$  nN) in contact mode and acquire AFM micrographs of all the nanografted nanoassemblages, that had undergone hybridization. This imaging section allow us to compare the AFM micrographs of DNA arrays to the same after hybridization with DNA-peptide conjugates and measure their respective relative heights.

### **Beta-2-microglobulin binding assay**

Starting with the equilibration of the macro-environment on the sample for binding assay, the sample was rinsed (twice) and incubated with phosphate buffer (binding buffer) for 10 minutes. Afterwards, the binding assay was carried out in contact with known concentration of  $\beta$ 2m solution, at 25°C for period of 1 hour within the Binder oven (model FD 115, Binder GmbH, tuttlingen, Germany). Subsequently, the grafting section (where cyclic peptide was immobilized) was imaged at low force ( $\sim 0.2$  nN) in contact mode of the AFM. AFM topographic height measurement using Igor Pro, 6.3A was employed to measure the relative height of DNA-cyclic peptide assemblage before and after binding assay with  $\beta$ 2m. This is the general procedure for DNA-pep461, DNA-pep331 and DNA-pep381 nanoassemblage.

### **AFM imaging of DNA-peptide assemblages before and after the $\beta$ 2m recognition assay**

At this point, prior to acquiring the topographic images of the DNA-peptide assemblages on the surface, it is expedient to perform tip-sample alignment using the SV-alignment software. SV-alignment software simply superimposes the reference tip-sample image on the video panel and the live tip-sample image. This alleviates the problem of loss of experimental domain, where the grafted patches were located on the sample, and support

continuous experiment on the same experimental domain. MFP-3D Bio was used simultaneously to obtain the topographic images of the double stranded DNA-peptide assemblages before and after the  $\beta$ 2m recognition assay. The topographic images were acquired taken in contact mode with soft probe (CSC38/ no Al) at the lowest force ( $\approx 0.2$ nN) within the DNA free buffer (1M NaCl, TE1X buffer). Conversely, for the topographic images after the  $\beta$ 2m recognition assay, the images were acquired in 25 mM sodium phosphate, 50 mM NaCl, and pH 7.

#### 4.5 References

1. Lavi, A. *et al.* Detection of peptide-binding sites on protein surfaces: The first step toward the modeling and targeting of peptide-mediated interactions. *Proteins Struct. Funct. Bioinforma.* **81**, 2096–2105 (2013).
2. Xu, S. & Liu, G. Y. Nanometer-scale fabrication by simultaneous nanoshaving and molecular self-assembly. *Langmuir* **13**, 127–129 (1997).
3. Niemeyer, C. M., Sano, T., Smith, C. L. & Cantor, C. R. Oligonucleotide-directed self-assembly of proteins : semisynthetic DNA-streptavidin hybrid molecules as connectors for the generation of macroscopic arrays and the construction of supramolecular bioconjugates. **22**, 5530–5539 (1994).
4. Soler, M. A. *et al.* Computational design of cyclic peptides for the customized oriented immobilization of globular proteins. *Phys. Chem. Chem. Phys.* (2016). doi:10.1039/C6CP07807A
5. Tuerk, C. & Gold, L. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* **249**, 505–10 (1990).
6. Smith, G. P. & Petrenko, V. A. Phage Display. *Chem. Rev.* **97**, 391–410 (1997).
7. Russo, A. *et al.* In silico generation of peptides by replica exchange monte carlo: Docking-based optimization of maltose-binding-protein ligands. *PLoS One* **10**, 1–16 (2015).
8. Fortuna, S. *et al.* Chelating effect in short polymers for the design of bidentate binders of increased affinity and selectivity. *Sci. Rep.* **5**, 15633 (2015).

9. Solulink. Antibody-Oligonucleotide All-In-One Conjugation Kit Technical Manual. *Solulink Man. A-9202-001*, 1–38 (2011).
10. Bano, F. *et al.* Toward multiprotein nanoarrays using nanografting and DNA directed immobilization of proteins. *Nano Lett.* **9**, 2614–2618 (2009).
11. Mirmomtaz, E. *et al.* Quantitative study of the effect of coverage on the hybridization efficiency of surface-bound DNA nanostructures. *Nano Lett.* **8**, 4134–4139 (2008).
12. Castronovo, M. *et al.* Two-dimensional enzyme diffusion in laterally confined DNA monolayers. *Nat. Commun.* **2**, 297 (2011).

## 5.0 Conclusions and outlook

In this thesis I have focused on two major areas: 1) understanding the recognition and the specificity of highly specific and stringent enzyme reactions within confined and highly dense DNA nanoreactors that is akin to the DNA packaging in both eukaryotic and prokaryotic cells; 2) experimental validation and determination of binding affinity of *in silico* based cyclic peptides towards the solvent exposed binding sites on the biomarkers that are over-expressed in ovarian cancer, such as Beta-2-microglobulin ( $\beta$ 2m). Our methodologies and discoveries in each section are summarized as follows.

In the case of enzymatic reactions within confined DNA nanoreactors, we utilized nanografting and spontaneous formation of self-assembled monolayer to form both laterally confined and highly dense DNA nanoreactors, which then serve as test-bed for enzymatic reactions. Using a novel method for the analysis of side-by-side AFM height profiling, our results showed that a stringent and highly specific enzyme, such as BamHI, can have an unprecedented mode of recognitions, leading to a digital behaviour of its efficiency within non-cognate DNA nanoreactors, which depends upon the inherent molecular density and its level of homogeneity within the DNA phase in the nanoreactor.

In the subsequent chapters, we reported a synergy approach that combined nanografting and DNA-directed immobilization, to construct arrays of *in silico* cyclic decapeptides, that possess high affinity and selectivity towards the protein of interest, which is  $\beta$ 2m in our case. Using a label-free AFM differential height measurements, our findings showed that the three cyclic peptides, pep461, pep331 and pep381, have micro-molar affinity towards the binding site on  $\beta$ 2m, with the latter two peptides leading with 7  $\mu$ M sensitivity to  $\beta$ 2m.

### 5.1 Future perspectives

The discoveries established in this thesis – namely about enzymatic reactions within confined and heterogeneous molecular structure and smart *in silico* based cyclic peptides as prospective and promising binders for surface bound arrays for protein and biomarker screening – all elucidate a number of additional experiments and research lines, which I will briefly elucidate in this section.

### **5.1.1 Outlook on enzymatic reactions within DNA nanoreactors**

We plan to further study the function of BamHI within micro-scale DNA nanostructures, made up of DNA-2 (DNA molecules with restriction site RS= 5' CGATCA') generated by spontaneous formation of self-assembled monolayer of pre-hybridized dsDNA. This is a crucial part in chapter 2 that was not completed because of limited time available for experimentation. This additional experiment will help us to fully understand the thresholds of density within which the anomalous behaviour of BamHI in DNA-2 molecules is dominant.

Furthermore, it would be interesting to carry out real-time enzymatic reactions, using high speed atomic force microscopy (HS-AFM)<sup>1-3</sup>. Considering, the fact that a single turnover endonuclease reaction process requires a few milliseconds, HS-AFM will allow us to capture the recognition and the dynamics of BamHI within density-dependent DNA nanoreactors. Additionally, we could also obtain information about the diffusion mechanism of BamHI into both nanoscale and micro-scale confined DNA nanoreactors, as there is no information in the literature on real-time diffusivity of DNA-binding proteins within three dimensional, highly dense nanostructures, to the best of my knowledge.

### **5.1.2 Future works on cyclic DNA-peptide arrays for biomarker detection**

Based on the promising results obtained in chapter 4 with respect to the two distinct cyclic peptides, pep331 and pep381, we plan to utilize AFM differential height measurements to obtain quantitatively and in a label-free way, the dissociation constant ( $K_D$ ) of each peptide and if possible determine the binding cooperativity within the DNA-peptide nanoassemblage towards each binding site on the  $\beta$ 2m.

Afterwards, we plan to couple the two peptides using a flexible linker, either through pegylation or DNA cross linking, and then immobilize the coupled peptides on surface using the combination of Nanografting and DNA-directed immobilization, while the surface-linking-DNA sequence will be different from those used for coupling to avoid base-crossing during the hybridization process. We believe this approach of coupling, which was inspired by chelation chemistry, will yield a higher binding affinity.

## 5.2 References

1. Ando, T. *et al.* High-speed atomic force microscopy coming of age. *Nanotechnology* **23**, 62001 (2012).
2. Casuso, I., Sens, P., Rico, F. & Scheuring, S. Experimental evidence for membrane-mediated protein-protein interaction. *Biophys. J.* **99**, L47-9 (2010).
3. Monachino, E., Spenkelink, L. M. & van Oijen, A. M. Watching cellular machinery in action, one molecule at a time. *J. Cell Biol.* 1–11 (2016).  
doi:10.1083/jcb.201610025

## **Appendix A. pH controlled assembly of DNA tiles\***

*In this chapter, all the AFM based characterization of DNA tiles were performed by me, which includes self-assembling of DNA tiles in time-pH dependent fashion, optimizing the approach for adsorbing DNA tiles on mica substrate following the self-assembling process, image processing and analysis of DNA tiles using Igor pro software.*

Here, we demonstrate a strategy to trigger and finely control the assembly of supramolecular DNA nanostructures with pH. Control is achieved via a rationally designed strand displacement circuit that responds to pH, and activates a downstream DNA tile self-assembly process. We observe that the DNA structures form under neutral/basic conditions, while the self-assembly process is suppressed under acidic conditions. The strategy presented here demonstrates a modular approach towards building systems capable of processing biochemical inputs and finely controlling the assembly of DNA-based nanostructures under isothermal conditions. In particular, the presented architecture is relevant for the development of complex DNA devices able to sense and respond to molecular markers associated with abnormal metabolism.

### **A.1 Introduction**

In Nature, several vital cellular tasks, such as the formation of the cell membrane or of stable host-guest complexes rely on thermodynamically driven molecular assembly processes based on relatively weak interactions.<sup>1</sup> The assembly of such complexes is usually finely controlled by a series of biological inputs and molecular cues.<sup>2</sup> Inspired by this observation, researchers in the field of supramolecular chemistry have exploited non-covalent interactions to achieve controlled self-assembly of synthetic moieties and to build complex nanostructures of defined geometries.<sup>1,3</sup> Because of its predictable base-pairing interactions and its low synthesis cost, DNA represents one of the best biomaterials to design and assemble complex structures with nanoscale features. Such structures have reached a

---

\* The experiments presented in this chapter was done in collaboration with Dr. Alessia Amodio and was published in September, 2016, in a form of communication in Journal of the American Chemical Society, 2016, *138* (39), pp 12735–12738.

level of complexity that would have been impossible to imagine 20 years ago: using synthetic DNA oligonucleotides we can now build 2D and 3D nano-scale objects with virtually arbitrary shape.<sup>4,5</sup> Input-responsive DNA assemblies, engineered to exhibit functional dynamic behaviours such as opening and closing or moving in response to biochemical inputs, have been also built.<sup>6</sup>

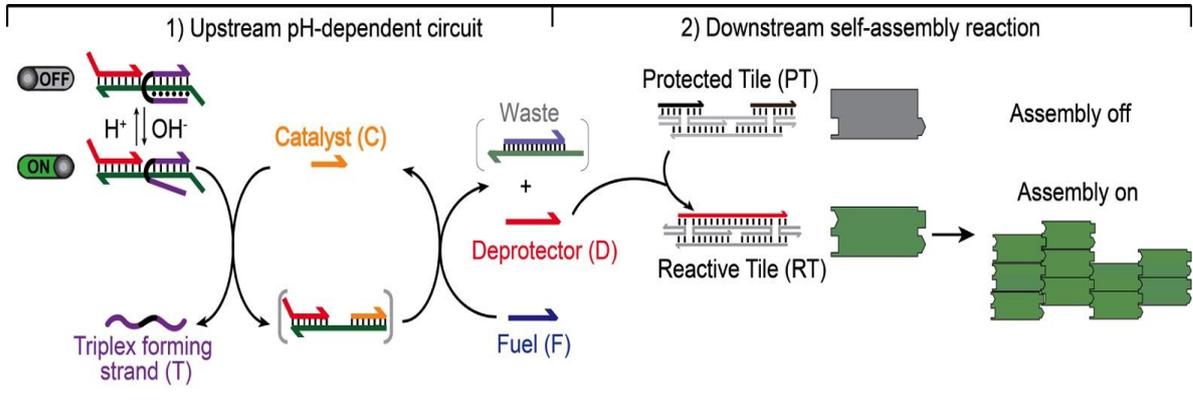
One of the current limitations of responsive DNA nanostructures is that responses are generally encoded in the structure itself (for instance, via input-responsive domains or aptamers), and can be modulated exclusively as a function of the input concentration or intensity. For example, DNA nanofabrication has been controlled directly using small synthetic ligands or light irradiation.<sup>7,8</sup> Finer control of assembly is desirable in many applications, such as sensing, metabolic engineering<sup>9</sup> or nanomanufacturing. This limitation can be mitigated by using input-triggerable DNA strand displacement circuits to control assembly.<sup>10</sup> Here, we demonstrate the viability of this approach with a pH triggerable DNA catalyst to control the assembly of DNA-based nanostructures.

## A.2 Results and discussion

We focus on DNA structures self-assembling from DNA tiles, one of the best-characterized approaches to build scalable DNA architectures.<sup>11,12</sup> In this strategy, DNA strands are designed to form rigid building blocks, called tiles (Fig. S11) that can self-assemble into lattices, ribbons, or tubular structures through single-stranded overhangs. DNA tile assembly can be triggered by an upstream DNA strand displacement circuit<sup>10</sup>: inactive (protected) DNA tiles can be activated (deprotected) by the output of a DNA catalytic circuit,<sup>13</sup> which is in turn triggered by a DNA catalyst molecule. This approach allows the isothermal assembly of DNA nanostructures, achieving control of the composition and timing of the process.<sup>14,15</sup>

To achieve pH-controlled, circuit-tunable assembly of DNA nanostructures we re-engineered a DNA strand displacement catalytic circuit (Fig. A.1) to be responsive to pH, and we used it to direct a well-known DNA tile self-assembly process based on double-crossover tiles (Fig. S11). pH-control is achieved by taking advantage of the well characterized pH sensitivity of triplex DNA, which requires the protonation of the N3 of cytosine in the third strand to form (average  $pK_a$  of cytosines in triplex structure is  $\approx 6.5$ ).<sup>16</sup> Specifically, we designed a pH-dependent substrate (Fig. A.1, purple strand) that prevents the strand displacement reaction with the Catalyst strand (C) at acidic pH, due to the formation of Hoogsteen interactions in addition to the Watson-Crick base-pairings. Only at

basic pHs (when Hoogsteen interactions are destabilized), the accessibility of the DNA substrate, needed for enabling such strand displacement circuit, is restored. In turn, tile assembly can occur efficiently upon Catalyst addition (Fig. A.1, right) under these pH conditions.

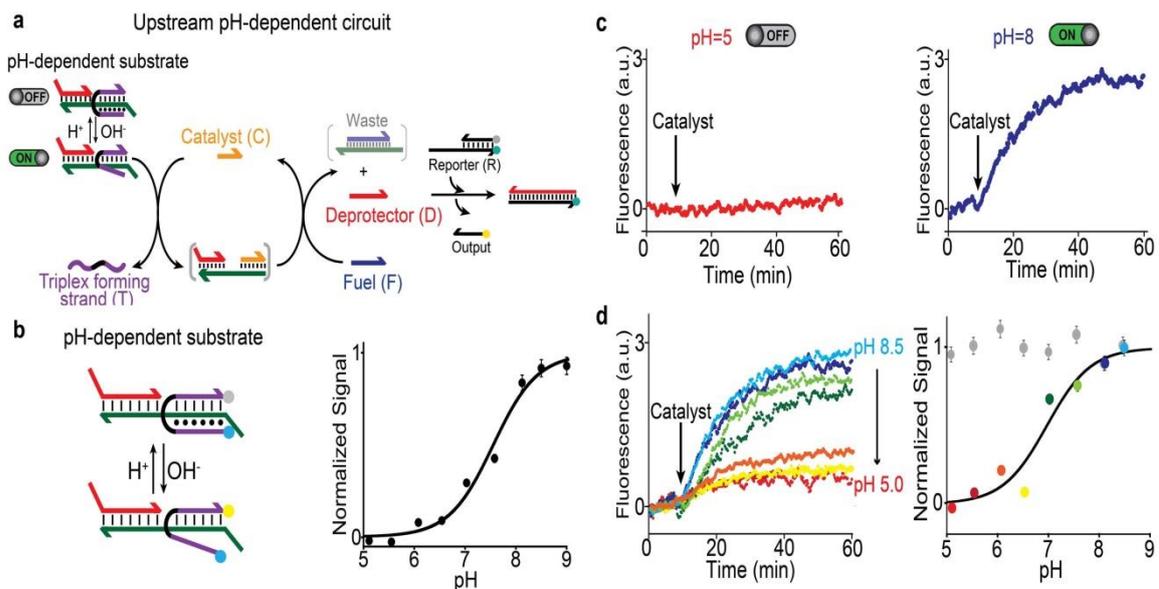


**Figure A.1:** pH-controlled self-assembly of DNA tiles. We have achieved pH-controlled DNA tile assembly by coupling an upstream re-engineered pH-controlled circuit with a downstream DNA tile self-assembly process. 1) In the pH-dependent upstream circuit a Catalyst (C) binds to a pH-dependent substrate leading to the release of a Deprotector strand (D). 2) The deprotector strand, in turn, activates a downstream self-assembly reaction by irreversibly associating with a protected tile (PT). This leads to reactive double crossover tiles (RT) self-assembling into lattices and nanotubes (right). The pH-dependent substrate is implemented with a clamp-like triplex-forming DNA strand that, under acidic pHs, can form a triplex complex inhibiting the strand displacement reaction with the Catalyst.

The designed upstream pH-dependent circuit can be finely controlled with pH. To demonstrate this, we have characterized in isolation the pH-dependent strand displacement circuit, by using an external, optically labelled reporter (R) that stoichiometrically reacts with the liberated Deprotector strand (D) (Fig. A.2a). We have first characterized the pH-dependency of the substrate complex by employing a triplex forming strand (T) (responsible for the substrate formation) labelled with a fluorophore and a quencher (Fig. A.2b, *left*). This allows us to monitor the folding and unfolding of the triplex structure at different pHs (Fig. A.2b, *right*). As expected, under acidic pHs (favouring triplex formation) a low fluorescence signal is observed, suggesting folding of the triplex complex. By increasing the solution's pH the fluorescence signal increases, consistently with the destabilization of Hoogsteen interactions (Fig. A.2a, *right*). pH of semi protonation (the average pKa due to

several interacting protonation sites) for this triplex complex is 7.5, which is in agreement with previous observations of similar triplex forming sequences.<sup>16,17</sup>

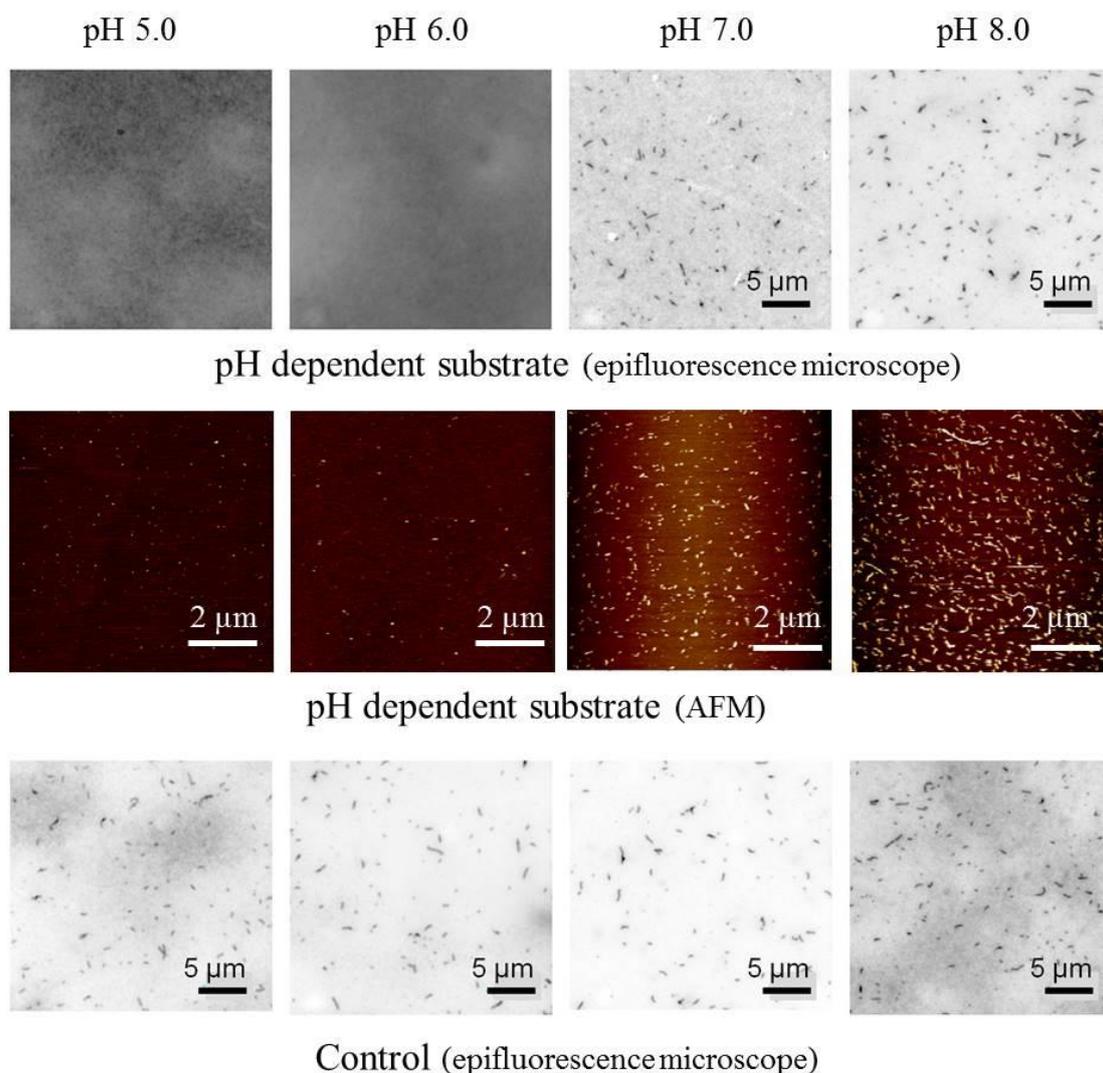
pH-dependent triplex formation in the substrate of the downstream catalytic network allows rational control of the output concentration by simply changing the solution's pH. At pH 5.0, which is acidic enough for having the clamp-like strand forming a triplex, inactive complex, the addition of the Catalyst strand (C) results in no significant fluorescence change (Fig. A.2c, *left*). This suggests that the circuit is fully suppressed. At pH 8.0, which inhibits triplex formation, strand displacement successfully proceeds with a fast kinetic upon catalyst's addition (Fig. A.2c, *right*). The activation level of the circuit can be modulated by changing the pH of the solution (Fig. A.2d and SI2) or the Catalyst concentration (Fig. SI3). However, probably due to the presence of the triplex forming sequence, the circuit catalytic efficiency is poorer than previously reported (Fig. SI4).<sup>10</sup> At high catalyst concentration, the complete circuit converts the gradual pH-dependency shown in Fig. A.2b into a digital-like response; in contrast, at low catalyst concentration the circuit response remains gradual. A control experiment involving a pH-independent substrate where the triplex forming portion has been substituted with a random sequence unable to form a triplex structure, shows no effect of pH over the entire pH range explored, and over a wide range of catalyst concentrations (Fig. A.2d, *right*, grey dots; Fig. SI5, SI6 and SI7).



**Figure A.2:** (a) Upstream pH-dependent DNA circuit. (b) Triplex formation in the pH-dependent substrate complex was studied by incorporating a pH-insensitive FRET pair at the ends of the clamp-like triplex forming strand (left), and measuring the fluorescence signal at different pHs (right). Stable triplex formation is observed only at pHs below 7.0. (c) The pH-dependent triplex complex

in the substrate inhibits strand displacement reaction upon catalyst addition (pH 5.0, left). At basic pHs the destabilization of Hoogsteen interactions leads to substrate activation, which allows strand displacement reaction in the presence of D (pH 8.0, right). (d) The pH-dependence of the catalyst/substrate reaction can be finely controlled at different pHs. Here, strand displacement reaction is followed by fluorescence measurements in a solution containing the pH-dependent substrate (10 nM), the fuel strand (F) (20 nM) and an external, optically-labelled reporter (30 nM) (see experimental section) that stoichiometrically reacts with the released deprotector strand (D) to give a fluorescence signal. The catalyst was added at a high concentration (30 nM) to better highlight pH-dependence of the circuit. All experiments were performed in a TAE 1x buffer + 15 mM MgCl<sub>2</sub>, at 25° C with the pH adjusted using small aliquots of HCl (1 M) and NaOH (1 M). Error bars here and in the following figures represent the average and standard deviations (average RSD% = 6%) of three independent measurements.

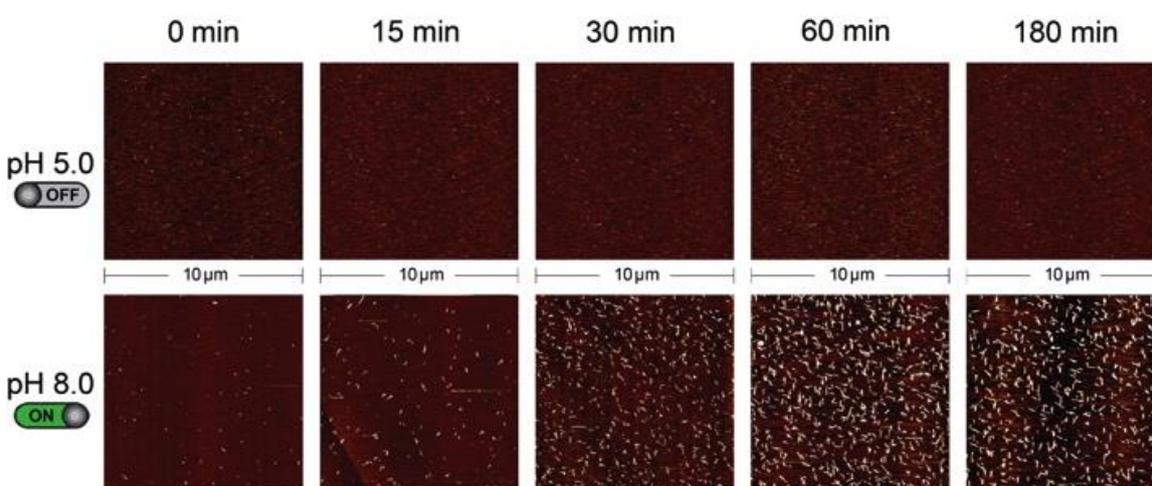
The pH controllable catalyst can be used to direct the assembly of DNA nanostructures using pH. To do this we have interconnected the above-characterized pH-dependent circuit with a DNA tile self-assembly process.<sup>10,11a,11b,12</sup> Fluorescence microscopy images (Fig. A.3, *top*) and AFM images (Fig. A.3, *centre*) confirm that tiles assemble only at neutral/basic pHs, while no assemblies are observed over the same reaction time at acidic pHs (pH 5.0 and 6.0). Tile assembly largely yields tubular structures according to fluorescence microscopy and AFM images. As a control experiment we used a pH-independent substrate and observed assembly of DNA tiles in the entire pH range investigated (Fig. A.3, *bottom*). Moreover, statistical analysis shows that the length and yield of nanotubes formed with the pH-dependent substrate under basic conditions (average length = 0.91 μm and yield = 23±10%, at pH 8.0) is comparable to those of the control pH-independent substrate at both pH 8.0 (length = 0.75 μm and yield = 12±6%) and 5.0 (length = 0.79 μm and yield = 19±12%) (for details see experimental section and Fig. SI8). As a further demonstration that pH does not affect the downstream tile assembly reaction, we have exogenously added the Deprotector to a solution containing protected tiles and observed pH-independent tile assembly (Fig. SI9). Conversely, the absence of Catalyst leads to no nanotubes formation (Fig. SI10)



**Figure A.3: pH-dependent self-assembly of DNA tiles.** (Top) The upstream, pH-dependent DNA circuit coupled with a downstream tile self-assembly process (Fig. A.1) allows to control DNA tile self-assembly with pH. At acidic pHs (pH 5.0, 6.0) no formation of assemblies is observed with optical fluorescence microscopy. By increasing the pH of the solution (pH 7.0, 8.0) we achieve evident formation of DNA lattices. (Centre) pH-dependent lattices were also imaged with atomic force microscopy (AFM). (Bottom) A control experiment using a pH-independent substrate (unable to form a pH-dependent triplex structure (Fig. SI5)) leads to pH-independent assembly of DNA tiles. All the experiments shown here and in the following figure were performed using the following concentrations of reagents: Protected Tile (PT) (200 nM), Fuel (F) (440 nM), pH dependent substrate or control substrate (220 nM) and Catalyst (C) (20 nM). The assembly was achieved in TAE 1x buffer + 15 mM MgCl<sub>2</sub>, at 25° C with the pH adjusted using small aliquots of HCl (1 M) and NaOH (1 M). For all the fluorescence microscopy experiments, a cy3-labeled tile central strand (t4, see SI) was used to detect nanotubes formation. AFM images of the pre-adsorbed nanostructures on freshly

cleaved mica (see SI) were obtained with AC mode in TAE 1x buffer + 15 mM MgCl<sub>2</sub> buffer, with 1 Hz scan rate, 256-pixel X 256-pixel image definition and processed with second order flattening.

The kinetics of DNA tile assembly was studied with AFM. Images derived from the reaction samples (upstream pH-dependent circuit + protected tiles) at different intervals after the addition of the Catalyst strand were obtained. No tiles assembly at pH 5.0 (Fig. A.4, *top*) and pH 6.0 (Fig. SI11) was observed during the entire time-frame investigated. At pH 8.0 a substantial number of assembled tiles can be observed within 30 minutes after Catalyst addition (Fig. A.4, *bottom*). We note that assembly of small lattices at pH 5.0 was observed under AFM after 4 days since the start of the reaction, presumably due to leak of the Deprotector from the substrate (Fig. SI12).<sup>10</sup>



**Figure A.4: Kinetic of pH-dependent self-assembly of DNA tiles.** The self-assembly of DNA tiles was followed by AFM images at different time intervals. No tile assembly is observed within the time-frame of the experiment at pH 5.0. In contrast, at basic pH (pH 8.0), DNA lattices can be clearly observed after just 15 minutes from Catalyst addition. See Fig A.3 for experimental details.

### A.3 Conclusion

In this work, we have demonstrated a modular architecture to regulate the self-assembly of DNA nanostructures with the sole change of pH. We did so by integrating an upstream pH-dependent strand displacement circuit into an already-characterized, downstream DNA tile self-assembly process. Our architecture is potentially relevant in biomedical applications of DNA nanotechnology. For example, pH dysregulation is a hallmark of several diseases,<sup>18</sup> including cancer. The availability of pH-triggerable DNA nanostructures will have intriguing potential applications in drug-delivery research.<sup>19</sup>

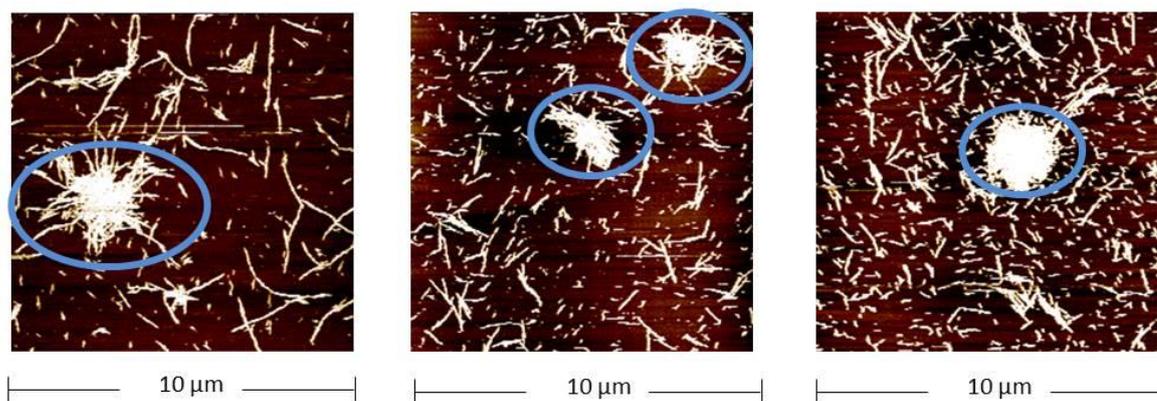
Moreover, because we have recently demonstrated that the pH-dependence of triplex DNA could be finely regulated by simply changing the triplex-forming sequence,<sup>16b</sup> we anticipate the possibility of programming the assembly of different DNA nanostructures at different pH thresholds. Finally, many enzymes are known that catalyse either proton-producing or proton-consuming reactions.<sup>20</sup> Therefore, different enzymes and enzymatic substrates could be used as molecular, functional inputs to control the isothermal assembly of DNA nanostructures.

The approach described here could also be expanded to consider a wider range of molecular inputs controlling the self-assembly of supramolecular structures. In fact, apart from pH, it is in principle possible to engineer upstream DNA strand displacement circuits activated by the presence of proteins, antibodies or other relevant biomolecular inputs so that the downstream self-assembly process is input-specific. In principle, logic or dynamic circuits could replace the catalytic network we considered, providing the system with expandable signal processing capacity. Finally, because strand displacement circuits can be designed to be highly specific and to respond orthogonally to a specific input, it would be possible to control the simultaneous assembly of different structures using multiple inputs in a highly programmable fashion, and thus leading to a better temporal and spatial regulation of the assembly processes.

#### **A.4 Future perspectives**

There are two possible future works that we plan to do with respect to the experiments on DNA tiles, especially in the AFM aspect of the work. Here they are;

- 1.) During the AFM-based characterization of DNA tiles, we observed specifically for control samples formed after 24-96 hours of self-assembling process, an association or clustering of DNA nanotubes into a weaving-mat like structure with few microns in size in x and y axis but with the same thickness in height (in z axis) (see Fig. A.5), which was not observed in the fluorescence experiments.



**Figure A.5:** pH independent DNA nanotubes weaved into DNA-mat revealed by the AFM characterization.

Based on this observation, we would like to understand the factors that are responsible for such formation and if this weaved-mat structure is facilitated by the substrate, concentration of the magnesium chloride in the self-assembling buffer or the association is catalysed by the small tiles in the solution. This information will help in developing DNA-based nano-substrate, which can be functionalized with different binders and then serve as a platform for diagnostics.

2.) Lastly, we observed that the DNA nanotubes either in control samples or pH-dependent samples at pH 8, were flat on the mica after been anchored on the surface, we believe this could be as a result of the surface interaction on the DNA nanotubes, the same for DNA nanotubes formed by Zhang *et.al*<sup>10</sup>. However, we plan to work on gold substrates and DNA tubes modified with thiol-derivatised, short DNA sequences, through which they can be anchored on the surface and hopefully allow us obtain hollow structure of DNA nanotubes as expected.

## A.5 Materials and methods

### Materials

Reagent-grade chemicals, including tris-base, glacial acetic acid, sodium chloride, magnesium chloride, ethylenediaminetetraacetic acid, ethanol and acetone (all from Sigma-Aldrich, St Louis, Missouri) were used without further purifications. HPLC purified oligonucleotides were purchased from IBA (Gottingen, Germany) and employed without further purification. The following oligos modified and non-modified were used:

**Protected tile**<sup>1-2</sup> (Fig. SII):

**t1:** 5'-ATACCATAGATCCTGATAGC-3'

**t2:** 5'-AGCAACCTGAAACCAGAATT-3'

**t3:** 5'-GAATTCTACTCGTGGATCTATGGTAT-3'

**t4:** 5'-AGAATTGCGTCGTGGTTGCTAGGTCTCGCTATCACCGATGTG-3'

**t4 (labeled):** cy3: 5'

AGAATTGCGTCGTGGTTGCTAGGTCTCGCTATCACCGATGTG-(cy3)-3'

**t5:** 5'-AATTCTGGTTTCACCTTAACGATACC-3'

**t6:** 5'-CGTTAAGGACGACGCAATTCTCACATCGGACGAGTAG-3'

**pH-dependent circuit (Fig 5.2):**

Fuel (F): 5'-AGCAACCTGAAACCACCTCTTTTCTTTCCC-3'

Catalyst (C): 5'-TTTTCTTTCCCTCACCATG-3'

Deprotector (D): 5'-ATAGATCCTGATAGCGAGACCTAGCAACCTGAAACCA-3'

Substrate (S): 5'-

CATGGTGAGGGAAAGAAAAGAGGGTGGTTTCAGGTTGCTAGGTCTC-3'

Clamp-like triplex forming strand (T): 5'-CCCTTTCTTTTCTCCC-GTTTG-

CCCTCTTTTCTTTCCC-3'

Duplex Control: 5'-ATCTTAACGTA CTGATTA-ATTCC-CCCTCTTTTCTTTCCC-3'

Reporter (Fig 5.2a):

R1: 5'-(cy3)-ATAGATCCTGATAGCGAGAC-3'

R2: 5'-TTGCTAGGTCTCGCTATCAGGATCTATR-(cy5)-3'

Labelled clamp-like triplex forming strand (strand forming complex T):

5'-(cy5)-CCCTTTCTTTTCTCCC-GTTTG-CCCTCTTTTCTTTCCC-(cy3)-3'

**Buffer conditions**

DNA oligonucleotides were suspended to a final concentration of 100  $\mu$ M and stored in 0.01 M TRIS + 0.01 M MgCl<sub>2</sub>, pH 7, at -20° C. In all experiments, we used solutions of TAE 1x + 15 mM MgCl<sub>2</sub> (starting pH = 8.2) with the pH adjusted with the addition of 1 M HCl or 1 M NaOH. All experiments were performed at 25° C.

### **Protected tile annealing**

The protected tile was prepared as reported elsewhere<sup>2</sup> with nominally correct stoichiometry at 5  $\mu$ M and annealed with a Bio-Rad Mastercycler Gradient thermocycler. The solution was brought down from 95° C to 20° C at a constant rate over a course of 6 h.

### **Substrates preparation**

All the complexes (Reporter, pH-dependent substrate and pH-independent substrate) used in the experiments were prepared by mixing the oligos necessary for the formation of the complex at equimolar concentration in TAE buffer 1x + 15 mM MgCl<sub>2</sub> pH 7 and were let to react overnight at room temperature.

### **Fluorescence measurements**

All fluorescence measurements were obtained using a Cary Eclipse Fluorimeter (Varian) with excitation at 548 nm ( $\pm$  5 nm) and emission at 563 nm ( $\pm$  5 nm). Strand displacement experiments were performed using a concentration of Reporter (R) of 30 nM, F of 20 nM and 10 nM of initial complex (pH-dependent substrate or pH-independent substrate). The Catalyst strand (C) was added at the selected concentration (from 1 nM to 100 nM) after 10 minutes to allow a stable baseline. The signal increase of Cy3 was followed for 1h after C addition.

### **pH-dependent self-assembly of DNA tiles**

This experiment was performed by adding (PT) 200 nM to a mixture reaction of (pH-dependent substrate or pH-independent substrate, the latter for Control experiment) (220 nM), F (440 nM), Catalyst was added at a concentration of 20 nM in TAE buffer 1x + 15 mM MgCl<sub>2</sub> at room temperature. Reaction time, 24 h.

### **Atomic force microscopy**

AFM topographic height images of the pH-dependent self-assembly nanotubes with 0 - 4 days' reaction time, were acquired using AC mode of MFP-3D Stand-Alone AFM (Oxford Instruments - Asylum research, Santa Barbara, CA, USA). Firstly, 20  $\mu$ l of 50  $\mu$ M of pH-dependent self-assembly nanostructures (with respect to reaction time) was deposited onto freshly cleaved mica substrate, fixed over a 150  $\mu$ l-volume, custom-made liquid cell. The adsorption step lasted for 20 minutes for all pHs. Afterwards, 100  $\mu$ l of DNA-free TAE/Mg buffer (TAE 1x with 15 mM Magnesium acetate tetrahydrate, with respect to the original

pH of the sample) was introduced onto the sample to have enough solution that covers the home-made liquid cell. The imaging parameters were as follow; AFM mode: AC mode in liquid, cantilever type: BL-AC40TS-C2 (a bioLever Mini Silicon tip on nitride lever), resonant frequency: 110 kHz (as specified by the manufacturer (Olympus, Japan)), spring constant: 0.09 N/m (as specified by the manufacturer), scan rate:1 Hz. We acquired 4 to 6 images, from different location of each substrate, at least 50  $\mu\text{m}$  apart with respect to each other. The topographic images were processed with 2nd order flattening, and analysed using Igor Pro 6.37 A.

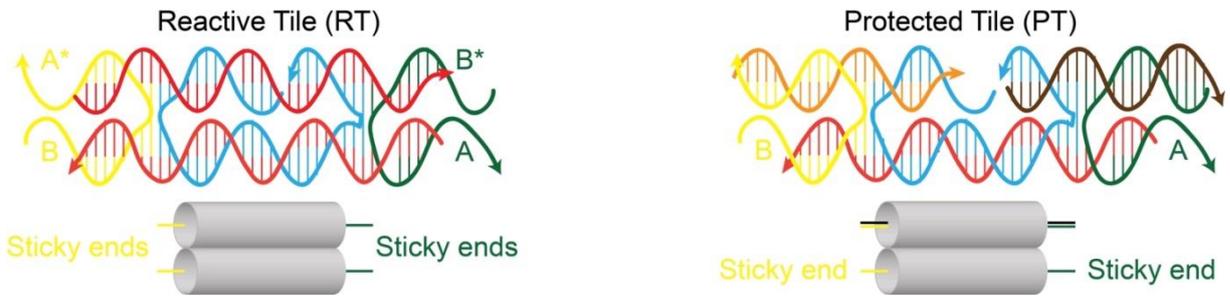
### **Fluorescence microscopy**

For fluorescence microscopy imaging the central strand of the tile (t4, see sequence above) was labelled at the 3' end with a cy3 fluorophore. An Axio Scope A1 ZEISS microscope was used. The emitted photons were collected by a 63x, air objective (ZEISS) and a monochrome CCD camera (AxioCam 503 mono - ZEISS). The images were analyzed using ZEN 2 lite (ZEISS) software. A 2  $\mu\text{L}$  drop of 50 nM of the sample was deposited between a clean microscope slide and a coverslip.

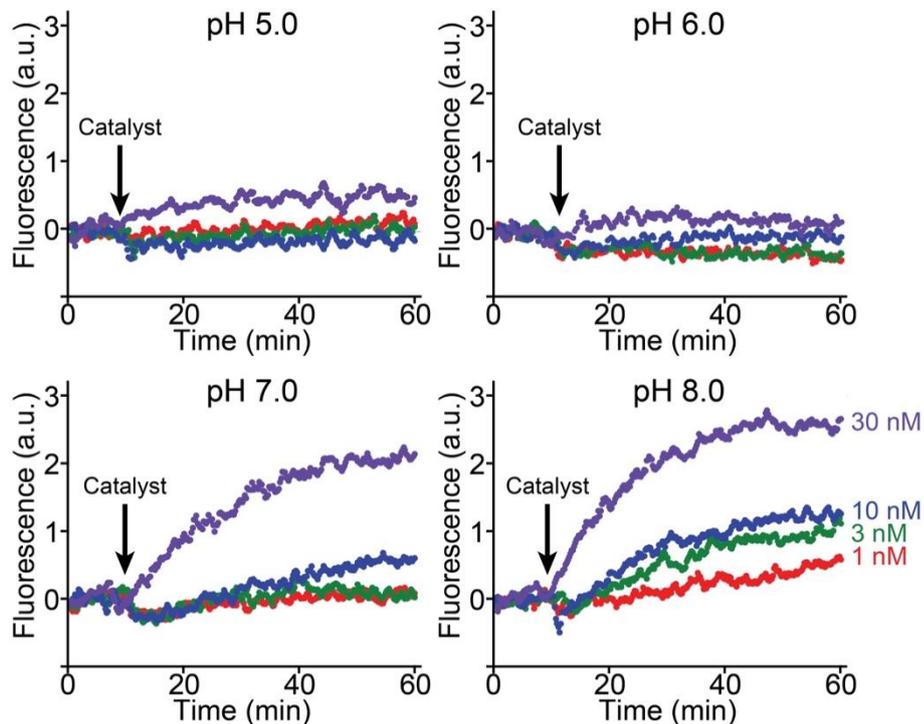
### **Measurement of tile-assembly yield**

In order to measure the yield of the tile assembly we applied the following procedure using the data obtained from fluorescence images (see Fig 5.3, SI8). By assuming an average thickness of the glass slide sample of 1.5  $\mu\text{m}$  and calculating the total area of imaging (3672  $\mu\text{m}^2$ ) we have calculated the volume of the imaged sample (7.3 pL). We have also assumed that the assembled tiles would equally deposit on the two sides of the glass slide. We have thus calculated the number of tiles from the length of the tubes (see statistical analysis in Fig. SI8) (each tile is 14.3 nm) and assuming that each tube has a fixed diameter of 7 tiles. In order to achieve the total yield of tile assembly we have compared the concentration of assembled tiles obtained with the above-described procedure to the concentration of total tiles in solution (200 nM). Of note, the concentration of total tiles in solution is assumed as equal to the concentration of the strands forming the tile (yield of tile formation from separated strands assumed as 100%).

## A.6 Supporting figures

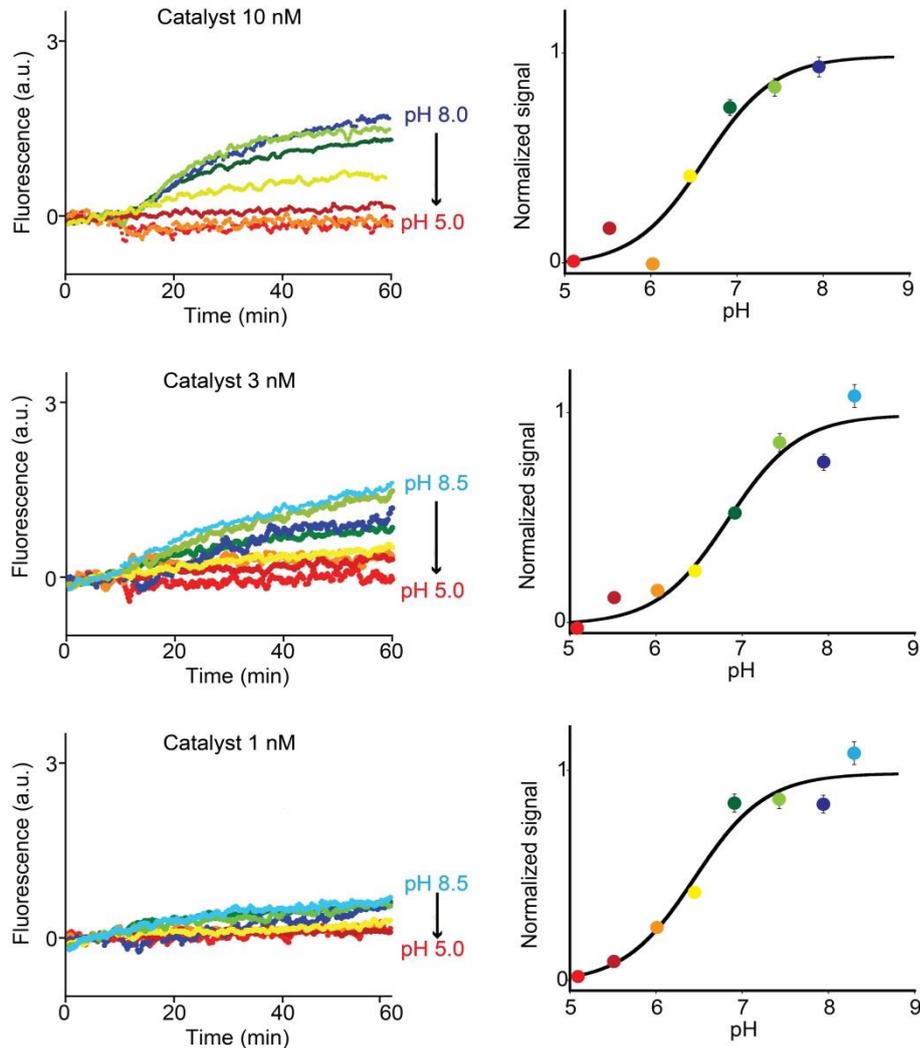


**Figure SI1: Reactive and protected tile.** (*left*) RT is made of five synthetic DNA strands that hybridize into a rigid rectangular core characterized by two double helices with a single-stranded five-base overhang (sticky end) at each corner (A, A\*, B, B\*). The presence of these sticky ends allows the self-assembly of RT. (*right*) In the PT, two of the four sticky ends (A\* and B\*) are protected with two complementary strands, thus leading to stable, monomeric tiles that cannot self-assemble<sup>1-2</sup>.

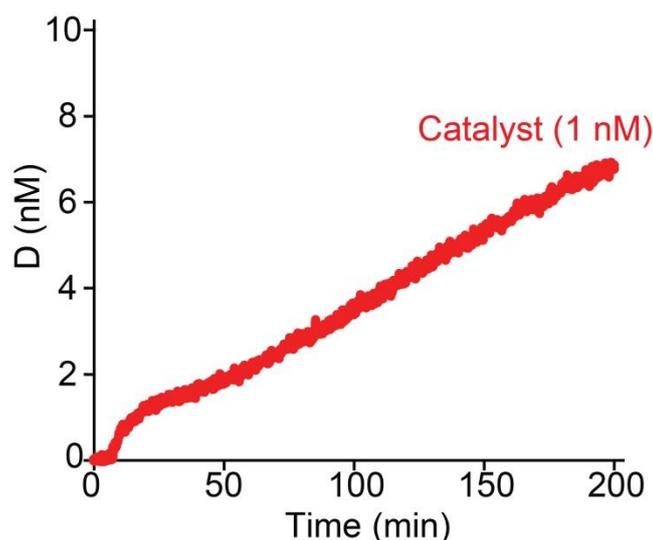


**Figure SI2.: Upstream pH-dependent strand displacement circuit characterization.** The addition of the Catalyst strand (from 1 nM to 30 nM) leads to no significant fluorescence signal change at pH 5.0 and 6.0, thus demonstrating that the formation of sequence specific Hoogsteen interaction at acidic pHs (Figure 2) strongly stabilize the pH-dependent substrate and inhibits the Catalyst-induced strand displacement reaction. In contrast, at neutral/basic pHs (pH 7.0 and 8.0) the destabilization of the Hoogsteen interactions allows successful strand displacement reaction upon

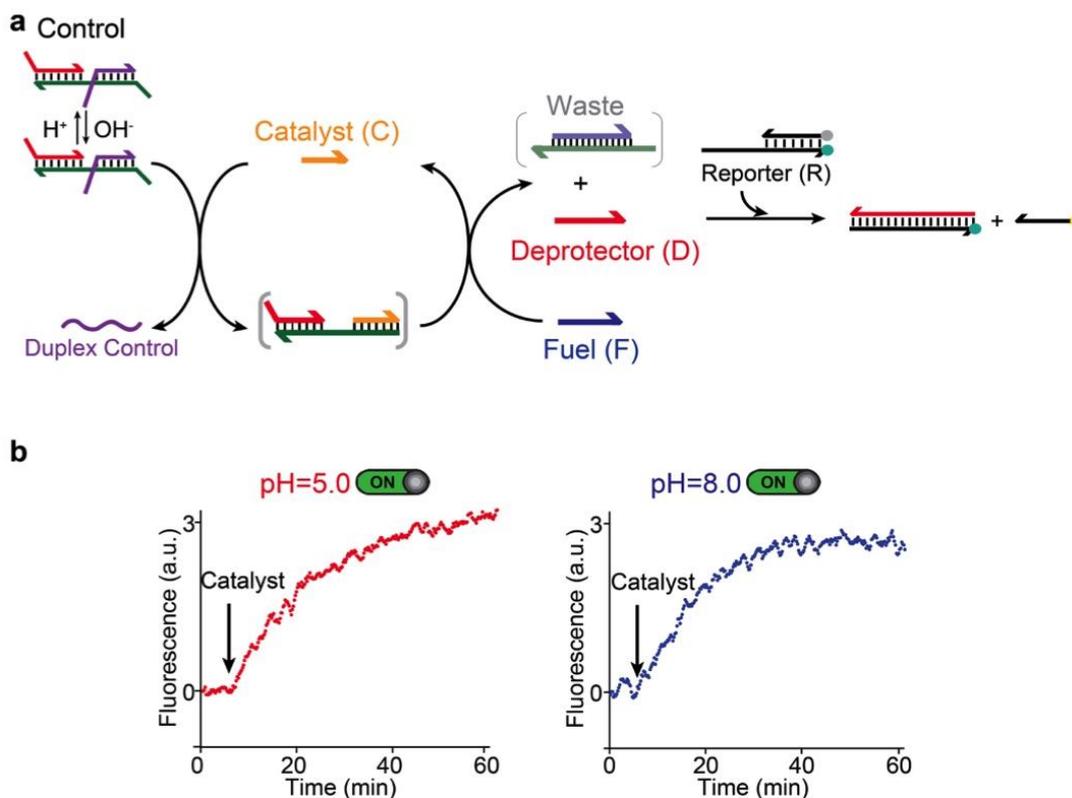
the addition of the Catalyst. The kinetic profiles suggest a faster kinetic at pH 8.0. Strand displacement is followed by measuring the fluorescence of a solution of complex pH-dependent substrate (10 nM) and Fuel (F) (20 nM) in the presence of a reporter (R) (30 nM), and after the addition of the Catalyst strand (S) at different concentrations in a TAE 1x buffer + 15 mM MgCl<sub>2</sub>, at 25° C.



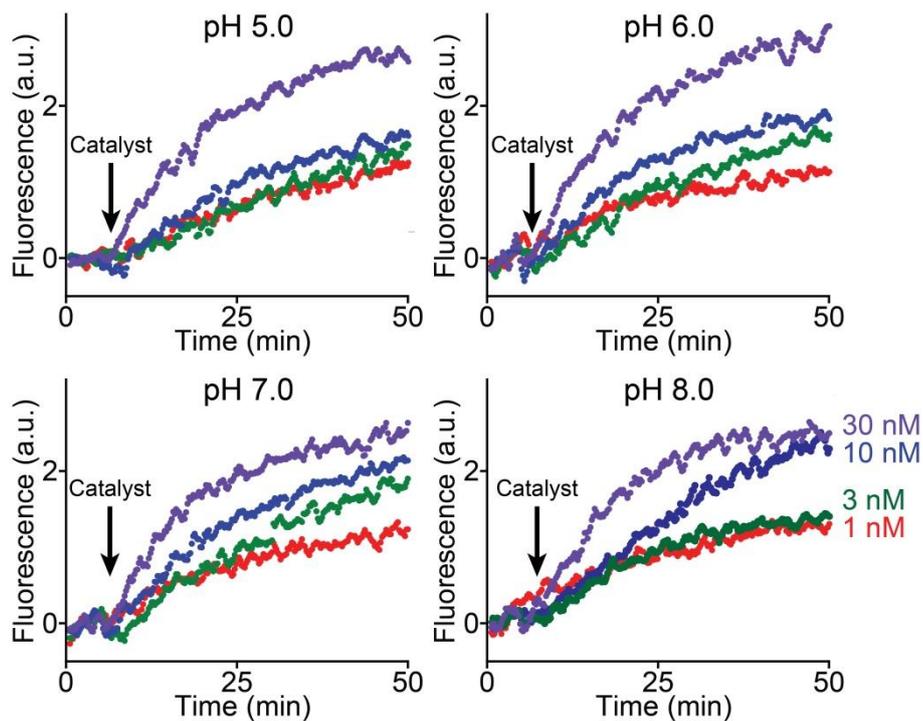
**Figure S13: Upstream pH-dependent strand displacement circuit characterization.** We can modulate the upstream strand displacement circuit using both pH and Catalyst concentration. At basic pHs the activation of the strand displacement process can be finely modulated using different concentrations of Catalyst (here are shown three different concentrations: 1, 3 and 10 nM). Strand displacement is followed by measuring the fluorescence of a solution of complex pH-dependent substrate (10 nM) and Fuel (F) (20 nM) in the presence of a reporter (R) (30 nM) (see also Figure 2a), and after the addition of the Catalyst strand (C) at different concentrations in a TAE 1x buffer + 15 mM MgCl<sub>2</sub>, at 25° C. The normalized signal shown in graph on the right are obtained by using the fluorescence value obtained at  $t = 60$  min in the kinetic experiments shown in the left. Standard deviation is within 5% for all data.



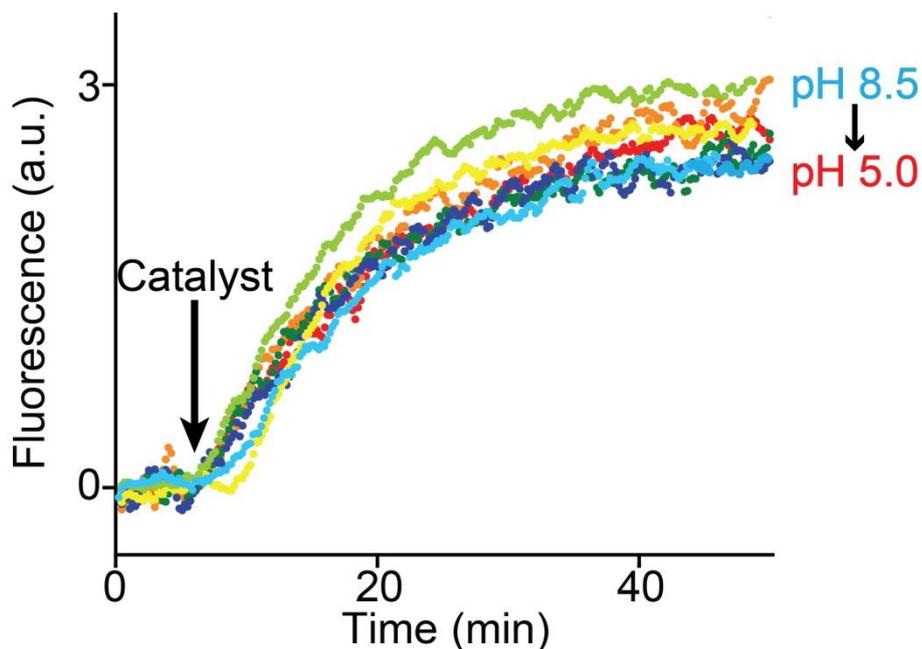
**Figure SI4: Characterization of the catalytic efficiency of the upstream circuit upon the addition of the Catalyst (C).** The addition of the Catalyst (1 nM) to a mixture containing complex pH-dependent substrate (10 nM) and Fuel (F) (20 nM) in the presence of a reporter (R) (30 nM) strand and under optimal pH conditions (pH 8.0) generates a catalytic response. The turnover number (moles of deprotector formed / moles of catalyst) of this catalytic circuit can be estimated by comparing the signal increase upon Catalyst addition (1 nM) with the signal expected from the addition of the deprotector (y-axis). The turnover number is 4 after 2 hours and 6 after 3 hours from Catalyst addition. This result demonstrates the catalytic efficiency of the upstream circuit to catalytically produce the deprotector strand upon the addition of the Catalyst. We note that the catalytic efficiency appears lower than that observed in the original paper by Zhang *et al.* (2) probably due to the effect of the triplex-forming tail on the strand-displacement reaction. The experiments were performed in a TAE 1x buffer + 15 mM MgCl<sub>2</sub> pH 8.0, at 25° C.



**Figure SI5: pH-independent strand displacement circuit (Control).** **a**) As a control system, we have designed a pH-independent substrate in which the triplex-forming portion of the pH-dependent substrate is substituted with a random tail that is unable to form a triplex structure. We have used an external reporter (R) that stoichiometrically reacts with the released Deprotector strand (D) to give a measurable signal. **b**) We observe an indistinguishable behaviour at both pH 5.0 and pH 8.0 after the addition of the Catalyst strand. Strand displacement is followed by measuring the fluorescence of a solution of complex pH-independent substrate (10 nM) and Fuel (F) (20 nM) in the presence of a reporter (R) (30 nM), and after the addition of the Catalyst strand (S) at a concentration of 30 nM in a TAE 1x buffer + 15 mM MgCl<sub>2</sub>, at 25° C.

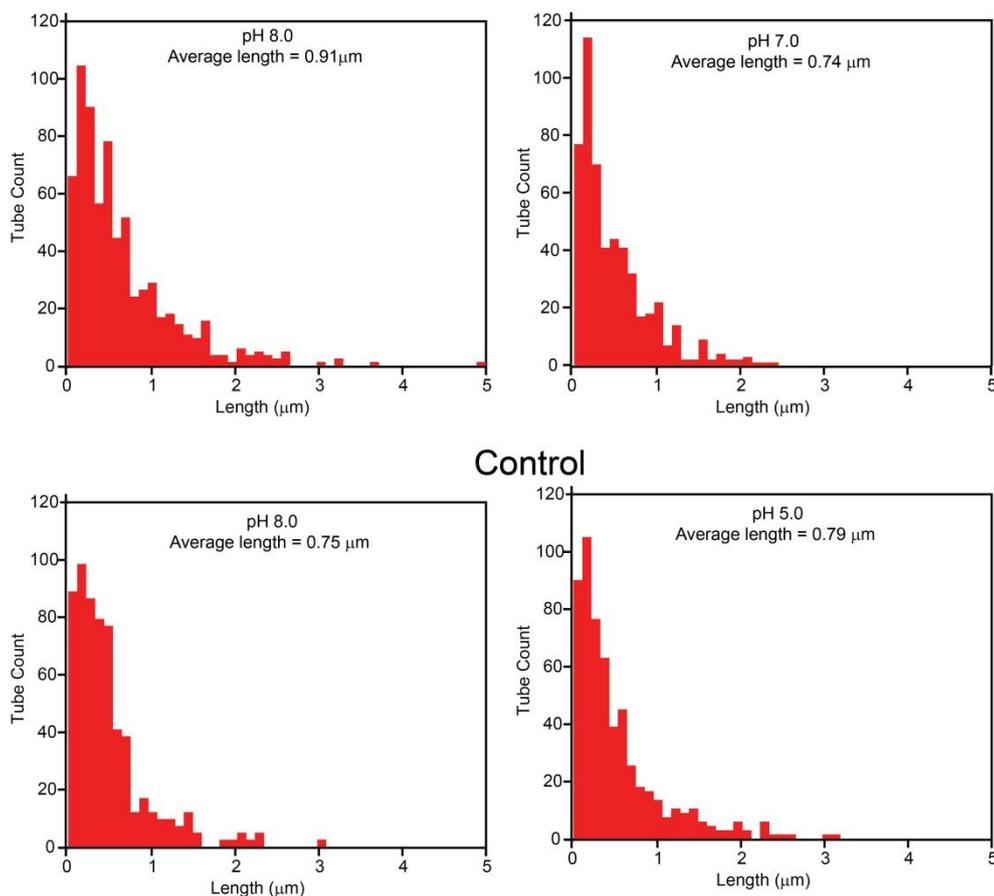


**Figure SI6: pH-independent strand displacement circuit (Control).** The control pH-independent strand displacement circuit (see Figure SI2) leads to successful release of the Deprotector strand at all the pHs investigated (pH 5.0, 6.0, 7.0, 8.0) with similar kinetic profiles. Strand displacement is followed by measuring the fluorescence of a solution of complex pH-independent substrate (10 nM) and Fuel (F) (20 nM) in the presence of a reporter (R) (30 nM) (see also Figure SI5), and after the addition of different concentrations of Catalyst strand (S) in a TAE 1x buffer + 15 mM MgCl<sub>2</sub>, at 25° C.

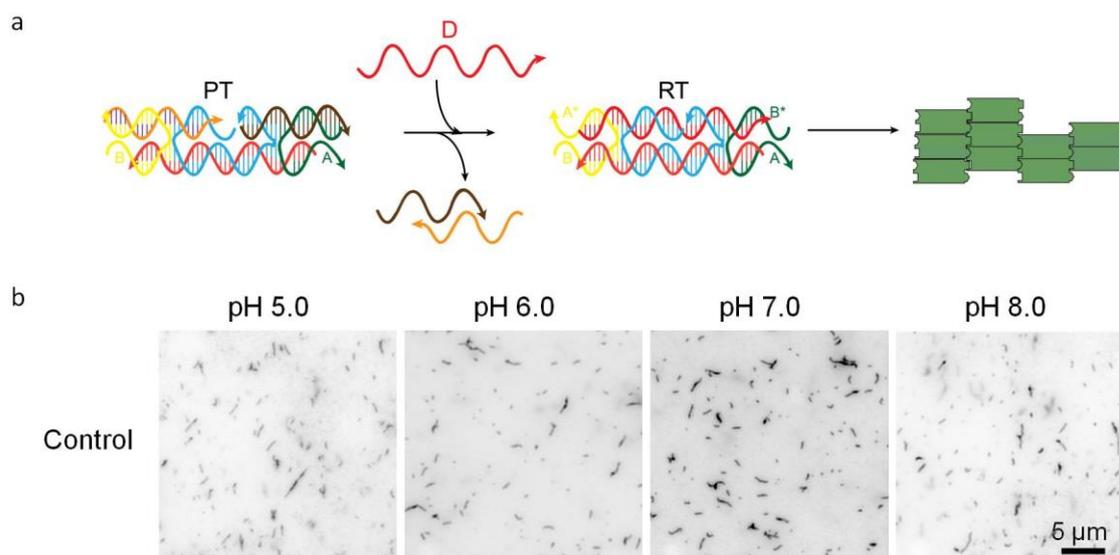


**Figure SI7: pH-independent strand displacement circuit (Control).** The control pH-independent strand displacement circuit (see Figure SI5) leads to successful release of the Deprotector strand at all the pHs investigated (from pH 8.5 (light blue) to pH 5.0 (red)) with similar kinetic profiles. Strand displacement is followed by measuring the fluorescence of a solution of complex pH-independent substrate (10 nM) and Fuel (F) (20 nM) in the presence of the reporter (R) (30 nM), and after the addition of the Catalyst strand (C) (30 nM) at different pHs in a TAE 1x buffer + 15 mM MgCl<sub>2</sub>, at 25° C.

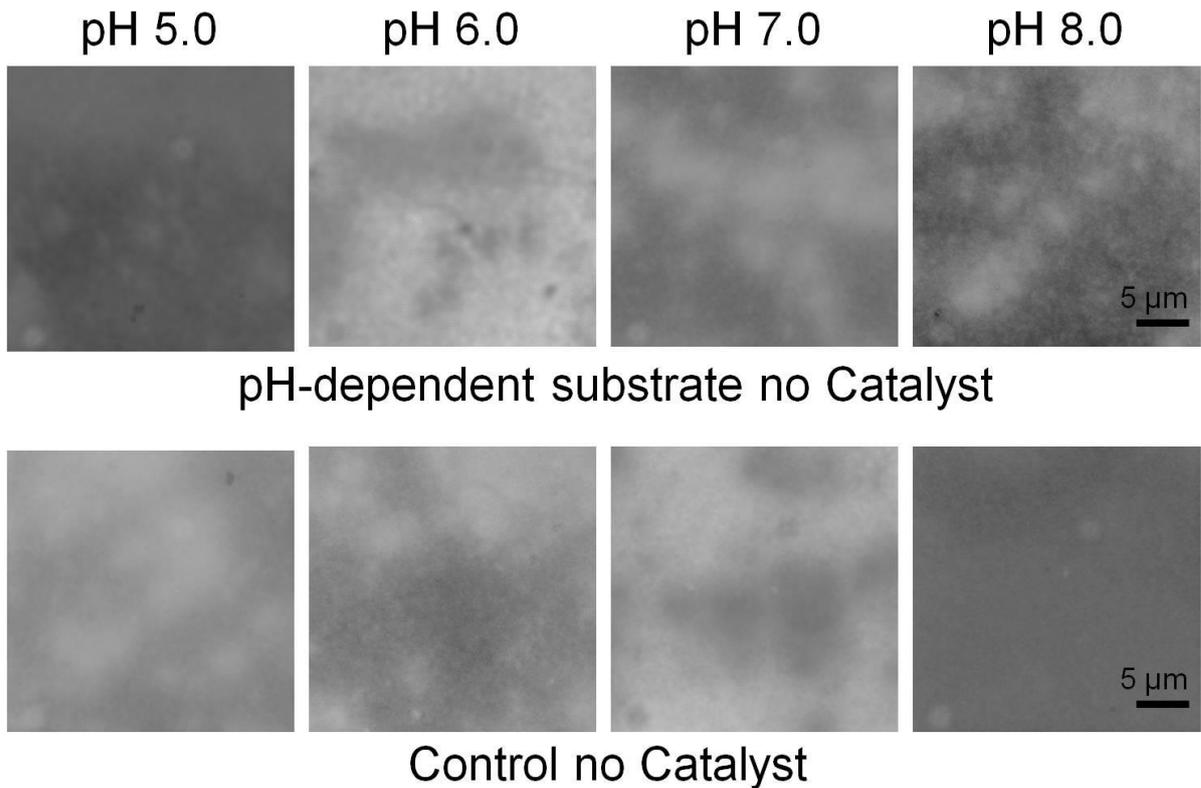
### pH-dependent substrate



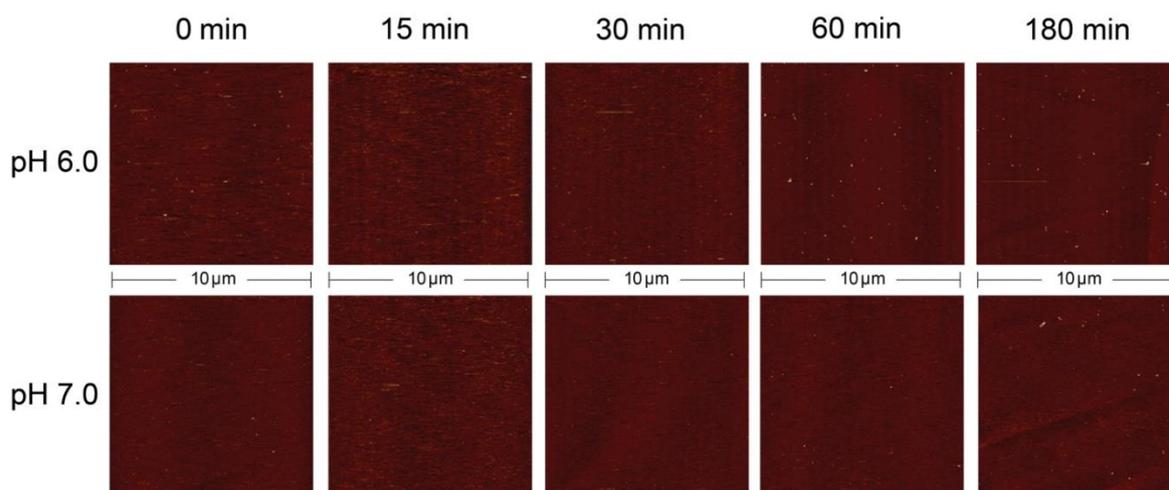
**Figure S18: Statistical analysis of assembled tiles length.** The average length of the assembled tiles has been evaluated using fluorescence microscopy images (12 images in total, 3 for each type of reaction). This experiment was performed by adding (PT) 200 nM to a mixture reaction of (pH-dependent substrate or pH-independent substrate, the latter for Control experiment) (220 nM), F (440 nM), Catalyst was added at a concentration of 20 nM in TAE buffer 1x + 15 mM MgCl<sub>2</sub> at room temperature. Reaction time = 24 h. The total number of assembled tiles analyzed is: pH-dependent substrate pH 8 = 579; pH-dependent substrate pH 7 = 526; Control pH 5 = 379 and Control pH 8 = 255.



**Figure SI9: Control of pH-dependence of DNA tile self-assembly reaction.** To demonstrate that DNA tile self-assemble with similar efficiency over the entire pH range tested, we have performed a control experiment in which the Deprotector strand (D) has been exogenously added to a solution of Protected Tiles (PT). **a)** The Deprotector (D) reacts with the Protected Tiles (PT) by displacing the two protector's strands and thus exposing the other two sticky ends. In turn, strand displacement leads to the formation of readily Reactive Tiles (RT) that isothermally form DNA nanotubes. **b)** Fluorescence microscopy of the above control experiment shows that self-assembly of DNA tiles occurs at all the pHs investigated (pH 5.0, 6.0, 7.0, 8.0). Experiments were performed by using a concentration of the Protected Tile of 200 nM and a concentration of the Deprotector strand of 220 nM in TAE 1x buffer + 15 mM MgCl<sub>2</sub>, at 25° C.

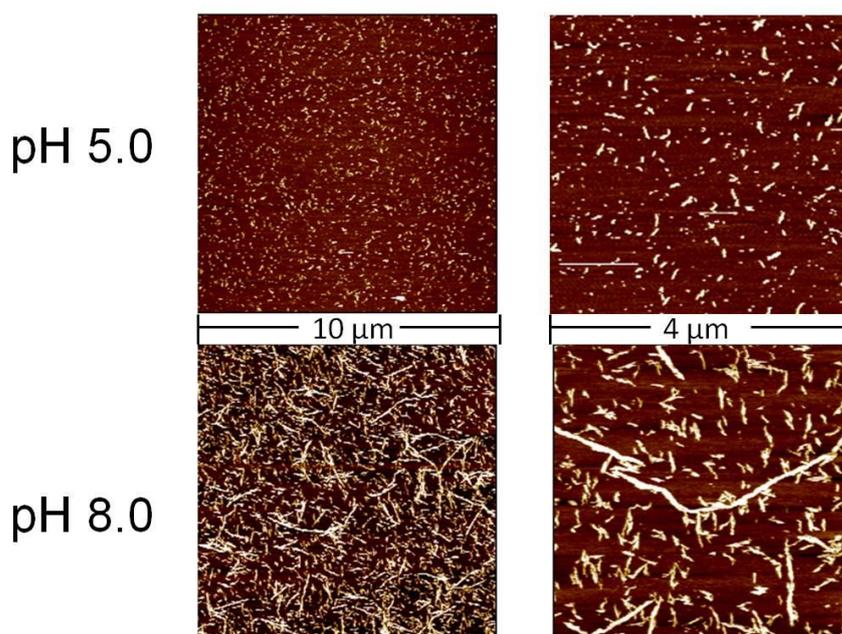


**Figure SII10: Fluorescence microscopy images of the integrated reaction network in the absence of Catalyst.** In the absence of the Catalyst strand no self-assembly of DNA tiles is observed for all the pHs investigated (pH 5.0, 6.0, 7.0, 8.0) using both the pH-dependent circuit and the pH-independent control circuit. All the experiments were performed with the same concentration of reagents: PT (200 nM), F (440 nM), pH-dependent substrate / pH-independent substrate (220 nM), in TAE 1x buffer + 15 mM MgCl<sub>2</sub>, at 25° C. For all the fluorescence microscope experiments, a cy3-modified version of the tile central strand (t4, see Material) was used.



**Figure SII1: Kinetic of the pH-dependent nanotubes self-assembly.** Kinetic of the integrated system at pH 6.0 (*first row*) and pH 7.0 (*second row*). All the experiments were performed with the same concentration of reagents: PT (200 nM), F (440 nM), pH-dependent substrate (220 nM) and C (20 nM), in TAE 1x buffer + 15 mM MgCl<sub>2</sub>, at 25° C.

### pH-dependent substrate



**Figure SII2: AFM characterization after 4 days of incubation.** The self-assembly of DNA tiles was evaluated by AFM images after 4 days of incubation. We note that assembly of small lattices at pH 5.0 (top row) is observed under AFM after 4 days since the start of the reaction, presumably due to leak of the Deprotector from the substrate<sup>1</sup>. At the same reaction time we observe DNA nanotubes formation at pH 8 (bottom row). Experiments were performed with the following concentrations of

reagents: PT (200 nM), F (440 nM), pH-dependent substrate (220 nM) and C (20 nM), in TAE 1x buffer + 15 mM MgCl<sub>2</sub>, at 25° C.

## A.7 References

1. Mattia, E.; Otto, S. *Nat Nanotechnol* **2015**, 10, 111.
2. (a) Desai, A.; Mitchison, T. J. *Cell Dev. Biol.* **1997**, 13, 83. (b) Lloyd, C.; Chan, J. *Nat. Rev. Mol. Cell Biol.* **2004**, 5, 13. (c) Li, R.; Gundersen, G. G. *Nat. Rev. Mol. Cell Biol.* **2008**, 9, 446.
3. (a) Li, J.; Nowak, P.; Otto, S. J. *Am. Chem. Soc.* **2013**, 135, 9222. (b) Subhabrata, M.; Fortuna, I.; Ferrante, C.; Scrimin, P.; Prins J. L. *Nat. Chem.* **2016**, 8, 725.
4. (a) He, Y.; Tian, Y.; Ribbe, A. E.; Mao, C. D. *J. Am. Chem. Soc.* **2006**, 128, 15978. (b) Ke, Y.; Ong, L. L.; Shih, W. M.; Yin, P. *Science* **2012**, 338, 1177. (c) Wei, B.; Dai, M.; Yin, P. *Nature* **2012**, 485, 623. (d) Zhao, Z.; Liu, Y.; Yan, H. *Nano Lett.* **2011**, 11, 2997.
5. (a) Fu, J.; Liu, M.; Liu, Y.; Woodbury, N. W.; Yan, H. *J. Am. Chem. Soc.* **2012**, 134, 5516. (b) Fu, J.; Yang, Y. R.; Johnson-Buck, A.; Liu, M.; Liu, Y.; Walter, N. G.; Woodbury, N. W.; Yan, H. *Nat. Nanotechnol.* **2014**, 9, 531.
6. (a) Douglas, S. M.; Bachelet, I.; Church, G. M. *Science* **2012**, 335, 831. (b) Gerling, T.; Wagenbauer, K. F.; Neuner, A. M.; Dietz, H. *Science* **2015**, 347, 1446.
7. (a) Liu, Z.; Li, Y.; Tian, C.; Mao, C. *Biomacromolecules* **2013**, 14, 1711. (b) McLaughlin, C. K.; Graham, D. H.; Sleiman, H. F. *Chem. Soc. Rev.* **2011**, 40, 5647. (c) Yang, Y.; Endo, M.; Hidaka, K.; Sugiyama, H. *J. Am. Chem. Soc.* **2012**, 134, 20645.
8. (a) Dohno, C.; Atsumi, H.; Nakatani, K. *Chem. Commun.* **2011**, 47, 3499. (b) Han, D.; Huang, J.; Zhu, Z.; Tuan, Q.; You, M.; Chen, Y.; Tan, W. *Chem. Commun.* **2011**, 47, 4670. (c) Yang, Y.; Endo, M.; Hidaka, K.; Sugiyama, H. *J. Am. Chem. Soc.* **2012**, 134, 20645.
9. Delebecque, C. J.; Lindner, A. B.; Silver, P. A.; Aldaye, F. A. *Science* **2011**, 333, 470.
10. Zhang, D. Y.; Hariadi, R. F.; Choi, H. M. T.; Winfree, E. *Nat. Commun.* **2013**, 4, 1965.

11. (a) Mao, C., LaBean, T. H., Reif, J. H. & Seeman, N. C. *Nature* **2000**, 407, 493. (b) Winfree, E., Liu, F., Wenzler, L. A.; Seeman, N. C. *Nature* **1998**, 394, 539. (c) Zhang, S.; Fu, T. J.; Seeman, N. C. *Biochemistry* **1993**, 32, 8062.
12. (a) Chen, H.; Weng, T.-W.; Riccitelli, M. M.; Cui, Y.; Irudayaraj, J.; Choi, J. H. J. *Am. Chem. Soc.* **2014**, 136, 6995. (b) Rothmund, P. W. K., Ekani-Nkodo, A., Papadakis, N., Fygenon, D. K.; Winfree, E. J. *Am. Chem. Soc.* **2004**, 126, 16344.
13. (a) Zhang, D. Y.; Seelig, G. *Nat. Chem.* **2011**, 3, 103. (b) Zhang, D. Y.; Turberfield, A. J.; Yurke, B.; Winfree, E. *Science* **2007**, 318, 1121.
14. (a) Bath, J.; Turberfield, A. J. *Nat. Nanotechnol.* **2007**, 2, 275. (b) Krishnan, Y.; Simmel, F. C. *Angew. Chem. Int. Ed.* **2011**, 50, 3124.
15. (a) Seelig, G.; Soloveichik, D.; Zhang, D. Y.; Winfree, E. *Science* **2006**, 314, 1585. (b) Turberfield, A. J.; Mitchell, J. C.; Yurke, B.; Mills, A. P., Jr.; Blakey, M. I.; Simmel, F. C. *Phys. Rev. Lett.* **2003**, 90, 118102/1. (c) Wang, F.; Lu, C.-H.; Willner, I. *Chem. Rev.* **2014**, 114, 2881.
16. (a) Amodio, A.; Bin, Z.; Porchetta, A.; Idili, A.; Castronovo, M.; Fan, C.; Ricci, F. *J Am Chem Soc.* **2014**, 136, 16469. (b) Idili, A.; Vallée-Bélisle, A.; Ricci, F. J. *Am. Chem. Soc.* **2014**, 136, 5836.
17. Porchetta, A.; Idili, A.; Vallée-Bélisle A.; Ricci, F. *Nano Lett.* **2015**, 15, 4467.
18. Webb, B. A.; Chimenti, M.; Jacobson, M. P.; Barber, D. L. *Nat. Rev. Cancer* **2011**, 11, 671.
19. (a) Sellner, S.; Kocabey, S.; Nekolla, K.; Krombach, F.; Liedl, T.; Rehberg, M. *Biomaterials* **2015**, 53, e453.
20. Del Grosso, E.; Dallaire, A. M.; Vallée-Bélisle, A.; Ricci, F. *Nano Lett.* **2015**, 15, 8407.

## **Appendix B: SPR, NMR, ITC characterization of the peptides for chapter 3**

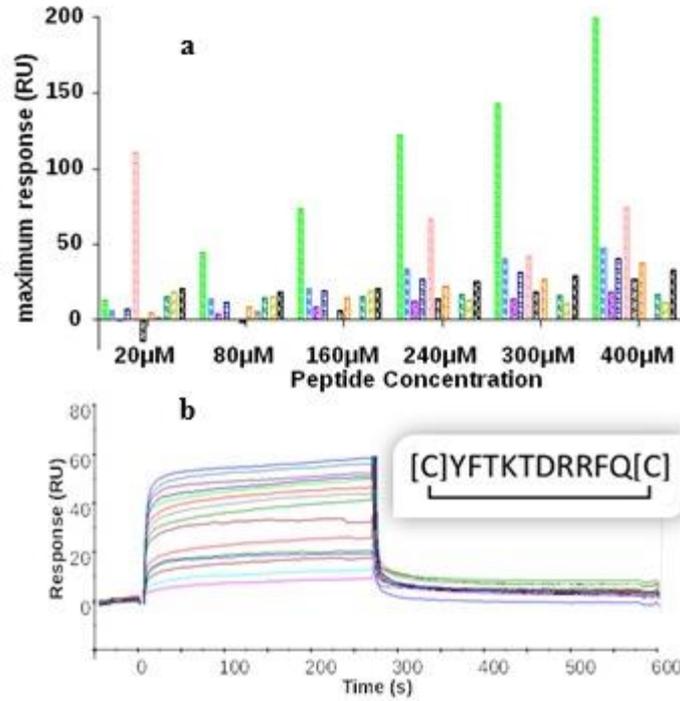
*By Dr Anna Russuo, Dr Daniella Marasco,*

*Ms. Cristina Cantarutti, Dr Alessandra Corazza*

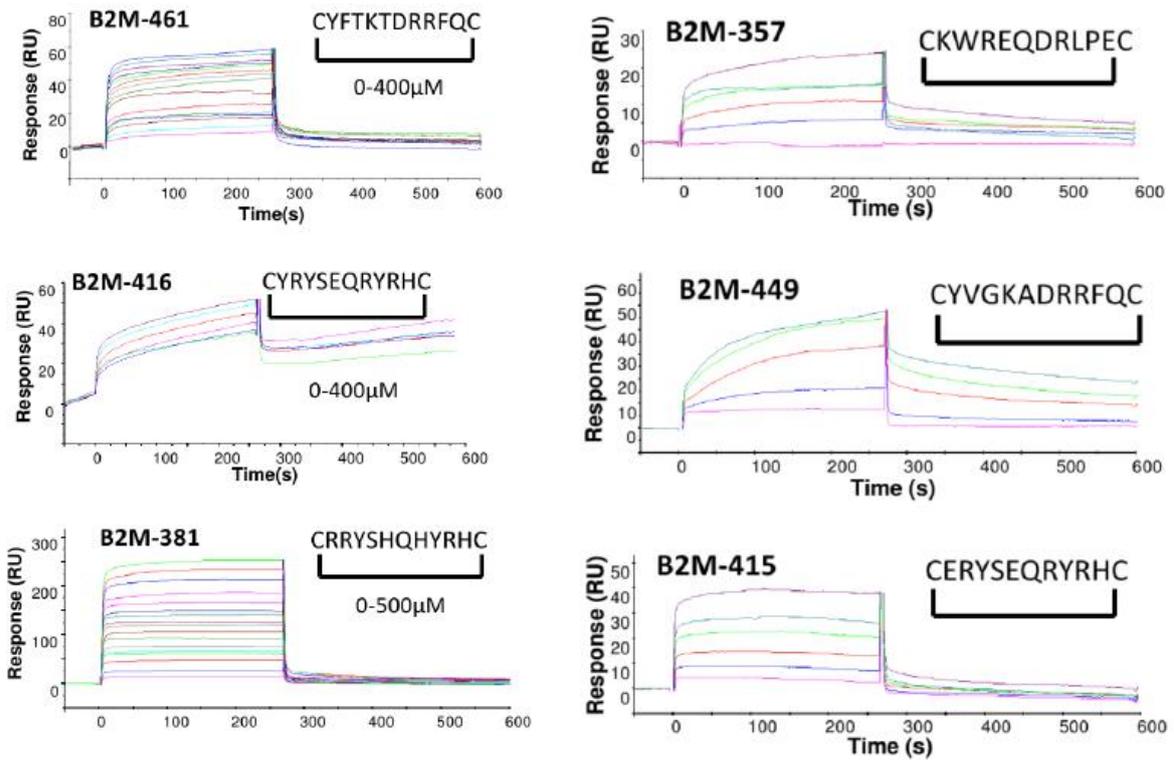
### **B.1: SPR results**

A secondary screening was carried besides the computer-based screening, by Dr Anna Russuo under the supervision of Dr Daniela Marasco (Department of pharmacy, University of Naples Federico II, Italy), using Surface Plasmon Resonance (SPR) and Isothermal titration calorimetry (ITC).  $\beta$ 2m was efficiently immobilized on a CM5 chip as described in the method section (see B.5), with an immobilization level of 3200 RU. The  $\beta$ 2m binding site did not contain lysine residues (Fig.3.0a) ensuring its right exposure during the amine-coupling/immobilization capture.

The eleven selected sequences were preliminarily assayed at six different concentrations of peptides (from 20 to 500  $\mu$ M). Histograms of response unit ( $RU_{max}$ ) for each concentration (see Fig. B.1a) show that six sequences gave dose response signals and were further analysed. The experimental curves corresponding to different concentrations of peptides (Fig. B.1b for peptide 461 and in Fig. B.2 for the other sequences) were fitted according to a single exponential binding model with 1:1 stoichiometry. This kinetic analysis led to the estimation of dissociation constant values reported in Table B.1, all in the micromolar range.



**Figure B1:** (a) Secondary Screening: SPR signals of eleven selected peptides at six different concentrations. (b) SPR sensogram of the selected sequence (sequence 461, in the inset).



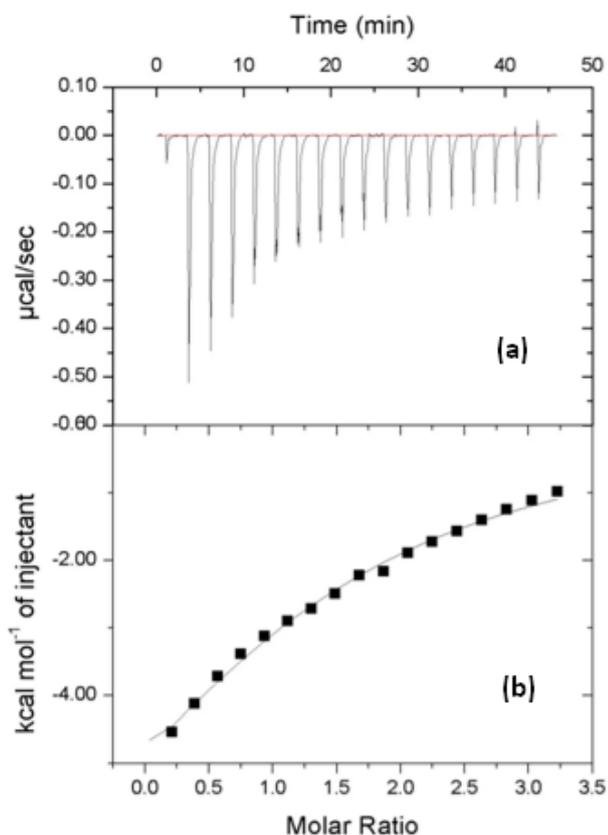
**Figure B2:** SPR sensogram for cyclic peptides that undergo secondary screening using SPR.

**Table B1.** Peptide sequences selected through SPR secondary screening, pI (isoelectric point) and dissociation constants toward  $\beta 2m$  are also reported ( $K_d$ ). Highlighted sequences (italics, red), tagged 461 and 381, are the peptides used for experimental validation by AFM in chapter 3 and 4 respectively.

<b>Id</b>	<b>Sequence</b>	<b>pI</b>	<b><math>K_d</math> (<math>\mu M</math>)</b>
357	[C]KWREQDRLPE[C]	6.17	$30 \pm 9$
<i>381</i>	<i>[C]RRYSHQHYRH[C]</i>	<i>9.50</i>	<i><math>38 \pm 9</math></i>
415	[C]ERYSEQRYRH[C]	8.06	$57 \pm 9$
<i>461</i>	<i>[C]YFTKTDRRFQ[C]</i>	<i>8.90</i>	<i><math>40 \pm 9</math></i>
449	[C]YVGKADRRFQ[C]	8.90	$36 \pm 8$
416	[C]ERYSEQRYRH[C]	8.86	$70 \pm 18$

## **B 2: ITC results**

Subsequently, pep461 was chosen out of the six peptides with a linear SPR response, because of its highest signal/noise response without containing His, this latter reduces possible problems of pH control during long MD simulations. Also, Pep461 reached saturation giving a  $K_d$  value of  $\sim 400 \mu M$ . By titrating aliquots of the same peptide into  $\beta 2m$  solution, the downward ITC titration peaks demonstrated the exothermic association between peptide and  $\beta 2m$  (Fig. B.3a). The ITC data were best fitted by a nonlinear least-squares approach to the ‘one set of sites’ binding model (see B.3b), providing a dissociation constant ( $K_d$ ) value of  $62 \mu M$  (Fig. B.3b), in agreement with the affinities showed for the kinetic based evolution of SPR assays. The ‘one set of sites’ model assumes a number of equally contributing binding sites having the same  $K_d$  and  $\Delta H$ . The fitted number of sites was  $1.34 \pm 0.14$ .

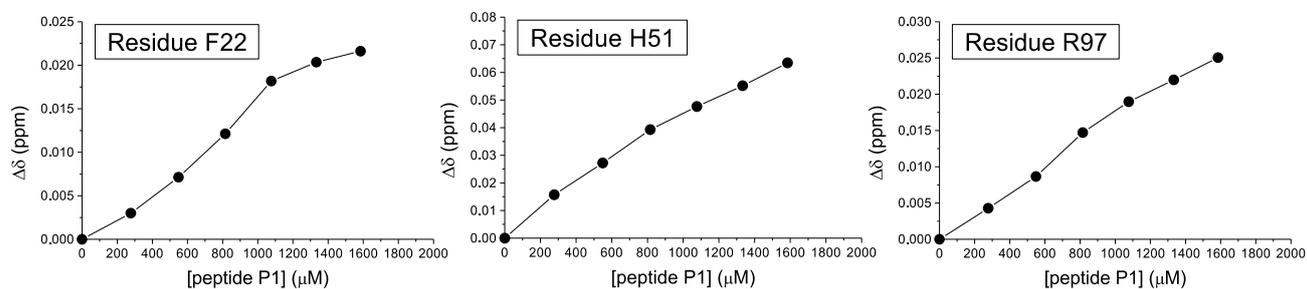


**Figure B3:** ITC Calorimetric (a) raw and (b) integrated data for  $\beta$ 2m (42  $\mu$ M) titration with peptide 461 (1 mM). Data fitting was achieved by ‘one set of sites’ model.

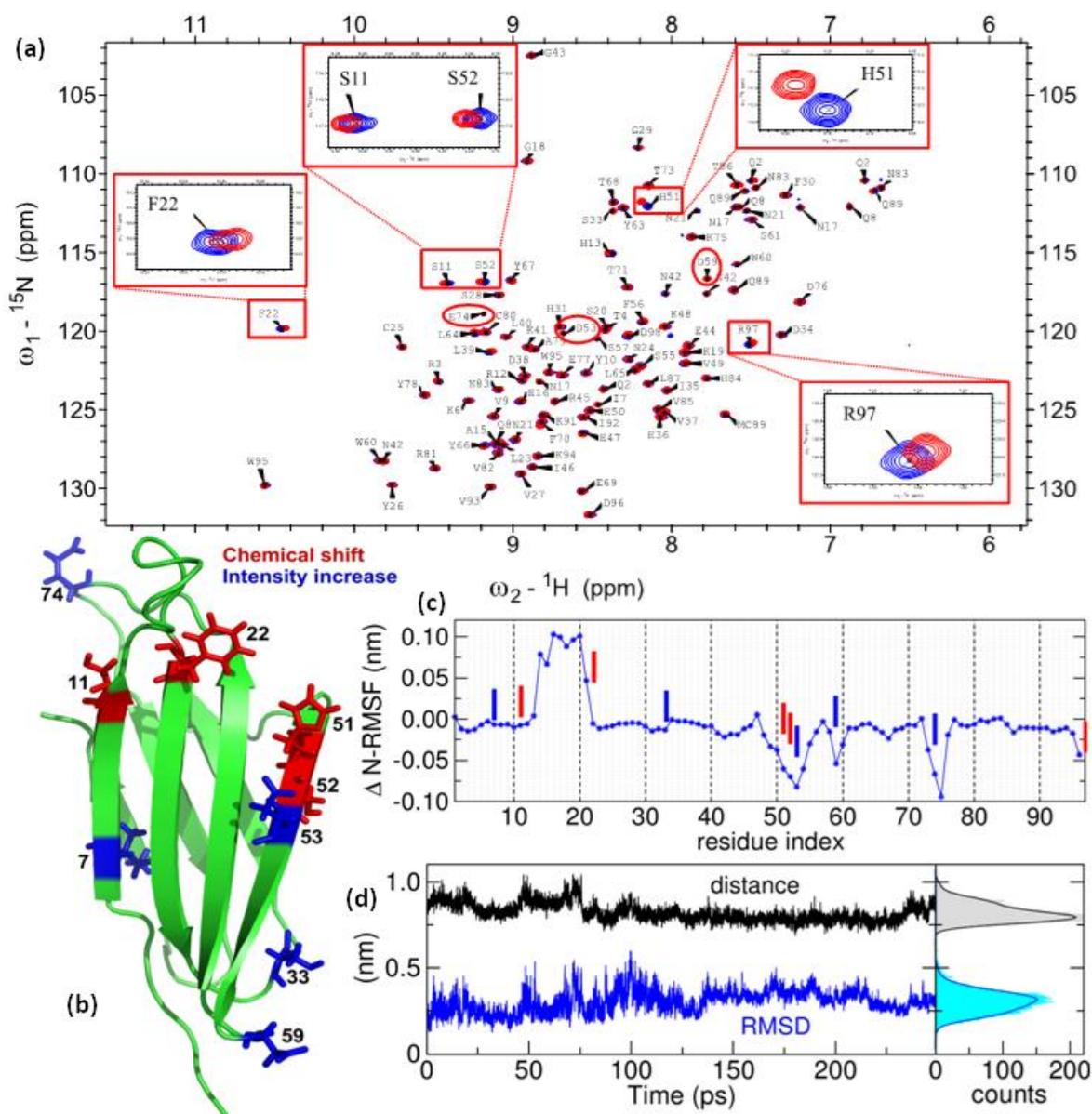
### B 3: NMR results

In addition to SPR and ITC results, Ms. Cristina Cantarutti performed NMR analysis under the supervision of Dr Alessandra Corazza, at the NMR facility, Department of medical and biological sciences, University of Udine. The NMR experiment is aimed to see whether the designed peptide was pinpointing the binding site it was designed for.  $^{15}\text{N}$  NMR was showing very minor chemical shifts in the 2D [ $^{15}\text{N}$ ,  $^1\text{H}$ ] HSQC spectra during the titration of  $^{15}\text{N}$  labeled  $\beta$ 2m with unlabeled 461 peptide (Fig. B.5a) and no saturation was reached in the chemical shift deviations plots for the shifted residues (Fig. B.4), an effect also observed in SPR (Fig. B.2) and affecting the binding affinity measured at saturation. Nevertheless, the NMR results showed that the affected amino acids were located on, or near the A, B, E, D  $\beta$ -sheets forming strands to which the putative binding site belongs (Fig. B.5b). When compared to the RMSF variation of the backbone amide nitrogen, the  $^{15}\text{NH}$  cross peaks of D53, D59 and E74 that grow in intensity and H51, S52, R97 that show a chemical shift deviation also display a N-RMSF variation upon binding (Fig. B.5c). Long MD simulations (250 ns, Fig. B.5d) revealed that the peptide 461 was highly mobile on the  $\beta$ 2m binding site

and could adopt several conformations. The peptide-protein distance calculated with respect to the  $\beta 2m$  binding site was distributed between 0.75 and 1.00 nm, with a distribution indicating multiple binding configurations located in the proximity of the chosen binding site (Fig. B.5d). Overall the peptide, while being highly mobile on the surface exposed site, tends to explore configurations close to the designed binding site.



**Figure B4:** Representative chemical shift deviations measured from 2D [ $^1\text{H}$ ,  $^{15}\text{N}$ ] HSQC experiments. Lines are only intended to guide the eye.

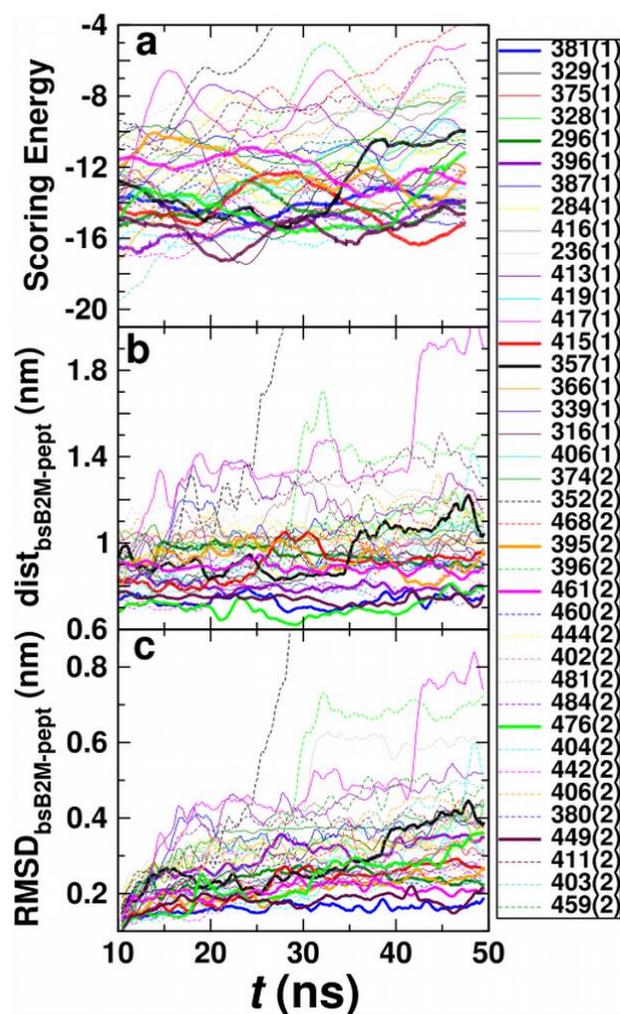


**Figure B5:** (a) Overlay of 2D [ $^1\text{H}$ ,  $^{15}\text{N}$ ] Heteronuclear Single-Quantum Correlation (HSQC) spectra of  $^{15}\text{N}$ -labeled  $\beta 2\text{m}$   $45 \mu\text{M}$  in the free form in blue and in the presence of peptide 461 in red (peptide/protein = 36.5/1). The peaks presenting the largest chemical shift deviations are boxed in red and enlarged in the excerpts. The signals affected by the most consistent intensity variations are highlighted with red circles. (b) NMR results depicted on  $\beta 2\text{m}$  structure: residues showing larger chemical shift deviations and intensity variations in the 2D [ $^1\text{H}$ ,  $^{15}\text{N}$ ] HSQC experiments are colored in red and in blue, respectively. (c) Nitrogen root mean square fluctuation difference ( $\Delta \text{N-RMSF}$ ) between free and peptide bound  $\beta 2\text{m}$  with vertical lines indicating amino acids presenting either chemical shift (red) or intensity increase (blue) upon binding, averages over 250 ns MD trajectories. (d) Peptide root mean square deviation (RMSD) at fixed  $\beta 2\text{m}$  binding site (blue) and distance between the peptide and its binding site (black) along the MD trajectory (left panel) and their distributions (right panel).

#### B 4: Additional results from computation-based screening

**Table B.2:** Scoring Energies (E, in kcal/mol) and sequences of the soluble peptides with the lowest Scoring E. The peptides are the output of two runs of the algorithm. Each run consists of three parallel replica exchange Monte Carlo optimisation run at the pseudotemperature T =0.3, 0.6, 0.9. The solubility has been predicted with Innovagen’s peptide calculator webtool [PepCalc.com. <http://pepcalc.com/>. Innovagen AB, 2012.]. Note the sequences used for experimental validation via AFM are highlighted in italics, red.

Index	Run 01			Run 02		
	Step (T)	Scoring E	Sequence	Step (T)	Scoring E	Sequence
1	<b>357</b> (0.6)	-25.4	CKWREQDRLPEC	374 (0.3)	-24.6	CYIGENDWRFHC
2	339 (0.3)	-25.3	CKWREQDRLSEC	352 (0.3)	-24.4	CYIGESDWRFHC
3	366 (0.6)	-25.0	CKWREFDRLPEC	468 (0.3)	-24.3	CYMTKSDRRFQC
4	329 (0.3)	-24.9	CVRYEWQHYRHC	<b>395</b> (0.6)	-24.3	CYRGQNDWRFHC
5	375 (0.3)	-24.7	CQRYSHQHYRHC	396 (0.6)	-24.2	CYRGQTDWRFHC
6	328 (0.3)	-24.6	CVRYEMQHYRHC	<b>461</b> (0.6)	<b>-24.1</b>	<b><i>CYFTKTDRRFQC</i></b>
7	432 (0.3)	-24.6	CVRYEEQWYRHC	460 (0.3)	-24.1	CYFTKSDRRFQC
8	<b>296</b> (0.3)	-24.5	CTRYFEQFYRHC	<b>444</b> (0.3)	-24.0	CYVIKADRRFQC
9	316 (0.3)	-24.5	CVRYERQHYRHC	402 (0.3)	-24.0	CYDGENDWRFHC
10	<b>381</b> (0.3)	<b>-24.4</b>	<b><i>CRRYSHQHYRHC</i></b>	481 (0.3)	-24.0	CYKFLDDWRFTC
11	406 (0.3)	-24.4	CYRYSEQHYRHC	484 (0.3)	-23.9	CYMTKVDIRFQC
12	<b>396</b> (0.3)	-24.3	CRRYSSQHYRHC	<b>476</b> (0.3)	-23.9	CYMTKSDIRFQC
13	387 (0.3)	-24.2	CRRYSGQHYRHC	404 (0.3)	-23.9	CYDGEADWRFHC
14	284 (0.3)	-24.2	CTRYILQKYRHC	460 (0.3)	-23.8	CYFGKSDRRFQC
15	<b>416</b> (0.3)	-24.1	CYRYSEQRYRHC	442 (0.6)	-23.8	CYTIKADRRFQC
16	236 (0.3)	-24.1	CTRYDWQHTRHC	406 (0.3)	-23.8	CYDGEADWRFSC
17	413 (0.3)	-24.0	CERYSEQHYRHC	380 (0.3)	-23.8	CYQGENDWRFHC
18	419 (0.3)	-24.0	CVRYSEQHYRHC	<b>449</b> (0.6)	-23.7	CYVGKADRRFQC
19	417 (0.3)	-23.9	CVRYSEQRYRHC	411 (0.3)	-23.7	CYRDKTDWRFHC
20	<b>415</b> (0.3)	-23.8	CERYSEQRYRHC	403 (0.3)	-23.6	CYRGKTDWRFHC



**Figure B6:** Primary screening: MD simulation results. Selected peptides are highlighted by thicker lines. (a) Average scoring energy, (b) Separation distance between  $\beta$ 2m binding site (BS- $\beta$ 2m) and peptide, and (c) RMSD of BS- $\beta$ 2m+peptide along the simulation time.

## B 5: Methods for computation-based design

### Peptides evolution

The  $\beta$ 2m structure with PDB code 1LDS<sup>1</sup> was first minimized with the steepest descent minimization method, with energy step size of 0.01 nm and a maximum force of convergence is  $< 1000$  KJ/mol/nm. The starting peptide was put close to the binding site, and the optimization started. At each optimization step a random mutation of one random peptide residue (excluding the extreme cysteine amino acids) was performed by using the AmberTool program<sup>2</sup>. The mutated structure was fully relaxed by three successive minimizations: (i) a partial minimization only for the side chain of the mutated residue, (ii) a partial minimization for the mutated amino acid and nearest neighbouring residues, and (iii) a global minimization. Each following REMD simulation consisted of 8 NVT replicas

with temperatures of 375, 391, 407, 423, 440, 458, 477 and 495 K, with all bond atoms constrained by using LINCS algorithm<sup>3</sup>, and with the backbone of the  $\beta$ 2m restrained by using the algorithm implemented in GROMACS (there isn't a particular name). Each replica was run for 1 ns, with a time step of 2 fs and an attempt of exchange every 2 ps. We clustered the peptide-protein samples obtained from all replicas by using the GROMOS method<sup>4</sup> as implemented in the gcluster program (part of the GROMACS package) with a cutoff of 0.105 nm. We discarded clusters containing less than 10 structures for the scoring evaluation.

### **Molecular dynamics simulations**

Each  $\beta$ 2m /peptide complex was first minimized using steepest descent minimization algorithm with 0.1nm step size to a maximum of 50000 steps and maximum convergence force of 1000.0 kJ mol<sup>-1</sup>nm<sup>-1</sup>. For NVT equilibrations were performed followed by one NPT using the leap-frog verlet integrator with the time step of 1 fs. The first equilibration to a total length of 25 ps was done by freezing both  $\beta$ 2m and peptide and the temperature of the previously minimized system was heated from 0 to 100 K (solvent plus ions) using velocity rescaling. The 2nd equilibration of 50 ps was done by keeping the temperature of solvent plus ions constant at 100 K and increasing the temperature of protein and peptide from 0 to 200 K. At the 3rd 50 ps equilibration we raised the temperature from 100 to 200 K and from 200 to 300 K for water plus ions and protein/peptide respectively. The fourth equilibration lasts 50 ps and the full system was equilibrated to 300 K. The fifth and last equilibration of 100 ps was done in NPT keeping the pressure constant to a reference value of 1 bar using the Parrinello-Rahman pressure coupling while the temperature remained constant to 300K. Production runs consists of 50 ns long NPT simulations with 2fs time step (plus additional 250 ns for the peptide 461). Configurations and energies were sampled every 10ps. All the interactions were characterized by amber99sb-ildn force fields<sup>5</sup> in the Periodic Boundary Conditions. The system was solvated using the tip4p water model<sup>6-7</sup>. The Particle Mesh Ewald summation accounted for long range electrostatic interactions. All the calculations as well as their analysis were performed as implemented in Gromacs-4.5.5 and Gromacs-4.6.2<sup>8</sup>. The binding energy of the peptide/protein complexes was estimated using Autodock Vina<sup>9</sup>.

## **B 6: Materials and methods for SPR, ITC, and NMR characterization**

### **SPR**

The interactions between the protein and computationally optimized peptides were measured using the BIAcore 3000 (GE Healthcare Milano, Italy).  $\beta$ 2m was immobilized at a concentration of 100  $\mu$ g/mL in 10 mM acetate buffer pH 5 (flow rate 5  $\mu$ L/min, time injection 7 min) on a CM5 Biacore sensor chip<sup>10</sup>, using EDC/NHS chemistry following the manufacturer's instructions. Residual reactive groups were deactivated by treatment with 1 M ethanolamine hydrochloride, pH 8.5. The reference channel was activated with EDC/NHS and deactivated with ethanolamine. The binding assays were carried out at 20  $\mu$ L/min at 25 °C, with 4.5 min contact-time. Peptides were diluted in the HBS running buffer (10 mM HEPES, 150 mM NaCl, 3 mM EDTA, pH 7.4). Analyte injections were performed at the indicated concentrations. The sensor surface was regenerated by using 10mM NaOH for 1 minute. The association phase ( $k_{on}$ ) was followed for 250s, whereas the dissociation phase ( $k_{off}$ ) was followed for 300 s. The instrument BIAevaluation analysis package (version 4.1, GE Healthcare, Milano, Italy) was used to subtract the signal of the reference channel.

### **ITC Isothermal titration calorimetry**

ITC experiments were carried out with an iTC200 calorimeter (Microcal/GE Healthcare). pep461 (1 mM) was titrated into a solution of  $\beta$ 2m protein (42  $\mu$ M). Data were fitted with 'one set of sites' model with ORIGIN software (GE Healthcare). Protein was dialysed against 10 mM phosphate buffer (pH 7.2) overnight, and all further dilutions of protein and peptide for ITC were made using the leftover external dialysate.

### **NMR characterization**

NMR experiments were recorded at 298 K on a Bruker Avance spectrometer operating at 500 MHz (<sup>1</sup>H). 1D <sup>1</sup>H spectra were acquired with 4096 data point, a spectral width of 16 ppm and 4096 scans. The water suppression was achieved by excitation sculpting scheme. 2D [<sup>1</sup>H, <sup>15</sup>N] HSQC spectra were acquired with 1024 and 128 points in the direct and indirect dimensions, respectively, and 400-1600 scans depending on the sample concentration, over spectral widths of 16 and 37 ppm in the <sup>1</sup>H and <sup>15</sup>N dimension, respectively. NMR samples were prepared in 25 mM sodium phosphate buffer and 50 mM NaCl at pH 6.95 and contained 8% D<sub>2</sub>O for lock purposes. The data were processed with Topspin 2.1 and analyzed with Sparky<sup>11</sup>. Chemical shift deviations were calculated as  $\Delta\delta$  (ppm) =  $[(\Delta\delta_H)^2 + (\Delta\delta_N/6.5)^2]^{1/2}$  where  $\Delta\delta_H$  and  $\Delta\delta_N$  are the chemical shift variations for <sup>1</sup>H and <sup>15</sup>N,

respectively<sup>12</sup>. The  $\beta$ 2m assignment was based on the file deposited on the Biological Magnetic Resonance Data Bank (Accession Code: 17165).

## B.7 References

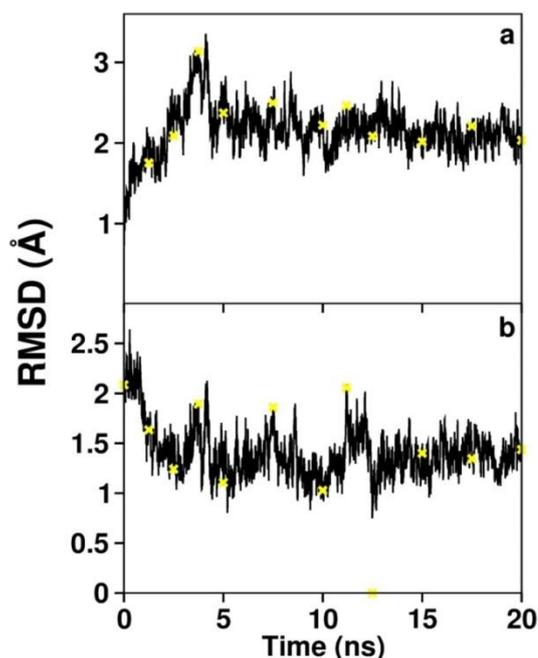
1. Trinh, C. H.; Smith, D. P.; Kalverda, A. P.; Phillips, S. E.; Radford, S. E. *Proceedings of the National Academy of Sciences* **2002**, 99, (15), 9771-9776.
2. Schafmeister, C.; Ross, W.; Romanovski, V. *University of California, San Francisco* **1995**.
3. Hess, B.; Bekker, H.; Berendsen, H. J.; Fraaije, J. G. *Journal of computational chemistry* **1997**, 18, (12), 1463-1472.
4. Daura, X.; Gademann, K.; Jaun, B.; Seebach, D.; van Gunsteren, W. F.; Mark, A. E. *Angewandte Chemie International Edition* **1999**, 38, (1-2), 236-240.
5. Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. *Proteins: Structure, Function, and Bioinformatics* **2010**, 78, (8), 1950-1958.
6. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *The Journal of chemical physics* **1983**, 79, (2), 926-935.
7. Jorgensen, W. L.; Madura, J. D. *Molecular Physics* **1985**, 56, (6), 1381-1392.
8. Hess, B.; Kutzner, C.; Van Der Spoel, D.; Lindahl, E. *Journal of Chemical Theory and Computation* **2008**, 4, (3), 435-447.
9. Trott, O.; Olson, A. J. *Journal of computational chemistry* **2010**, 31, (2), 455-461
10. Sreejit, G.; Ahmed, A.; Parveen, N.; Jha, V.; Valluri, V. L.; Ghosh, S.; Mukhopadhyay, S. *Plos Pathog* **2014**, 10, (10).
11. Goddard, T.; Kneller, D., San Francisco: University of California. In SPARKY.
12. Mulder, F. A.; Schipper, D.; Bott, R.; Boelens, R. *Journal of molecular biology* **1999**, 292, (1), 111-123.

## Appendix C: Computational details for the generation and screening of the peptides for chapter 4

by Dr Miguel A. Soler and Dr Sara Fortuna

### C.1 Binding site identification

First, we ran a 20 ns NPT molecular dynamics simulation of  $\beta$ 2m in explicit water at 300 K, after the respective equilibration process (100 ps NVT + 100 ps NPT). We exported conformations every 10 ps obtaining a set of 2000  $\beta$ 2m conformations. Then we calculated the root mean square deviation (RMSD) of the  $\beta$ 2m conformations taking as reference geometries  $\beta$ 2m structures at 0 and 12.5 ns (see Fig. C1). We selected structures every 2.5 ns and also those with higher values of RMSD, e.g. at 3.75 ns and 11.2 ns, forming a set of 12 structures. We note that the region A is the one chosen for us for the design of peptides as binders of  $\beta$ 2m.



**Figure C1:** Evolution of RMSD of  $\beta$ 2m taking as a reference (a) the 0 ns and (b) 12.5 ns structures.

Next, we calculate the putative binding sites of the set of  $\beta$ 2m structures using the webtool Peptimap (<http://peptimap.bu.edu/>). The Peptimap approach is based on the Fast Fourier transform-based docking protocol with statistical potential<sup>1</sup>, applied already to ligand binding site identification<sup>2</sup> and drug-able protein-protein interaction sites<sup>3</sup>. The tool

Peptimap has been tested<sup>4</sup> in a benchmark set of 21 peptide-protein complexes, identifying the binding site with 90% of success. We uploaded each pdb file of the  $\beta$ 2m conformation to the webtool and we activated the PPI mode, a minor modification to the scoring function which reduces the weight of a cavity term and thus optimize the results on a PPI test set. As output, we obtained for each  $\beta$ 2m conformation a pymol file with 6 different binding sites, ranked by the number of non-bonded contacts, and the residues of  $\beta$ 2m involved in each one of them. In order to classify the different binding site locations, we label 6 regions (A, B, C, D, E and F) on the  $\beta$ 2m (see Fig. 4.1(far right)).

**Table C1:** Solvent Accessible Surface Area (SASA) of each most probable binding site and sequences of residues involved in the most probable binding sites of every region.

Binding Site	SASA ( $\text{\AA}^2$ )	Sequence
A	772.5	Y11, Y27, D54, L55, S56, Y64, L65, L66
B	892.3	S34, D35, I36, E37, V38, R46, S53, D54, L55, L65, Y67
C	617.5	Q9, V10, I8, V94, K95, W96, D97
D	671.5	K42, I47, K49, T69, E70, F71, T72, Y79
E	521.1	R4, P6, F31, H32, P33, D35, I36
F	432.5	E37, V38, D39, R82

Additionally, we display in the Table C1 the Solvent Accessible Surface Area (SASA) of the most probable binding sites, calculated with pymol tool, and the sequences of the residues involved in each binding site. These results show that the binding site in the region B has the biggest surface area and involves the highest number of residues of  $\beta$ 2m, while the binding site in region A is the second one. We should note that the SASA of our initial 12-aa polyalanine peptide is around  $1500\text{\AA}^2$  and its value will increase until around  $2150\text{\AA}^2$  after the peptide optimization, so that the size of the binding site seems to be an important factor to consider before carrying out the peptide design simulations, principally in the case that we are considering pocket topologies. Except the binding sites in regions A and F, onto the surface of strands, binding sites are allocated in regions between two beta-strands (regions B and C), or surrounded by random coil chains of the  $\beta$ 2m (regions D and E). Therefore, it seems only the binding site in the region B is able to bind a 12-aa peptide, since it is the broadest.

In summary, we choose the binding site in the region B as a novel binding site for a second peptide, i.e. pep331.

## C.2 Peptide screening: Simulation details

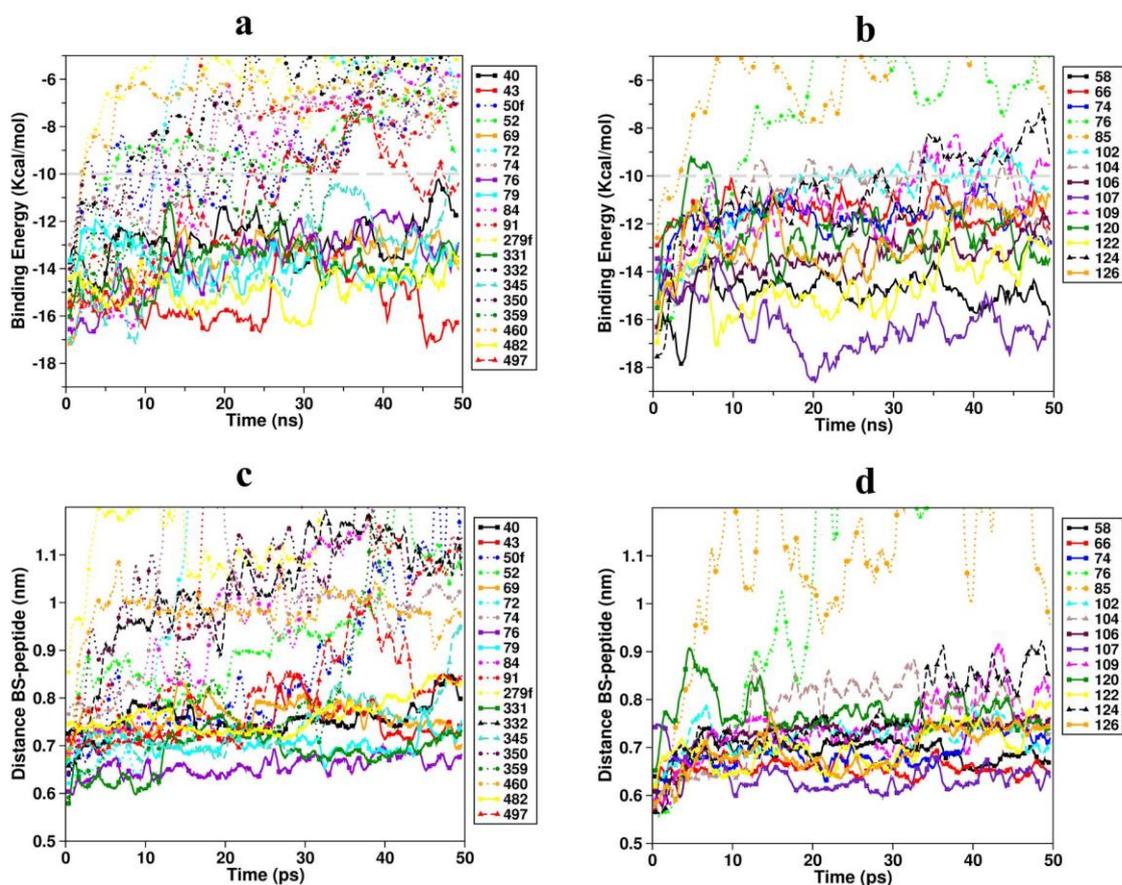
The computational screening of the designed peptides was carried out by means of Molecular Dynamics (MD) simulations. The equilibration protocol of the system before the production run is divided in 5 stages, 4 NVT and 1 NPT, in order to heat the system gradually, avoiding unrealistic high perturbations to the binding of the complex; first, we employed NVT equilibration of 25 ps, with  $\beta$ 2m+peptide frozen, and water and ions at 100 K; followed by NVT equilibration of 50 ps, with  $\beta$ 2m+peptide at 200 K, and the water and ions at 100 K ; and then NVT equilibration of 50 ps, with  $\beta$ 2m+peptide at 300 K, and the water and ions at 200 K. Subsequently, we used NVT equilibration of 100 ps, with the system at 300 K, followed by NPT equilibration of 200 ps, with the system at 300 K and 1 atm. We ran 50 ns NPT MD simulations for each  $\beta$ 2m-peptide complex in water with 0.15 M of NaCl, using amber99sb-ildn force field along with the tip3p water model, all implemented in Gromacs-4.5.5 molecular package<sup>5</sup>. We used a time step of 1 fs, constraining the h-bonds of the system. The temperature and the pressure established are 300 K and 1 atm.

$\beta$ 2m-peptide configurations were sampled every 250 ps. The 200 structures were scored by Vina<sup>6</sup>. Additionally, we analyze the h-bond and hydrophobic contacts of each  $\beta$ 2m-peptide complex at 0, 25 and 50 ns using LIGPLOT tool<sup>7</sup> to identify the residues of the binding site (BS) in every complex. Besides Vina binding energy, all descriptors used in this report were calculated using the tools implemented in Gromacs-4.5.5 molecular package<sup>5</sup>.

## C.3 Peptides screening: Computational results

In more detail, Fig. C2a-b shows the evolution of binding energies of the  $\beta$ 2m-peptide complexes as calculated by VINA. According to previous tests, Vina energies higher than -10 kcal/mol correspond to weak interactions between the protein and the peptide. In Fig. C2a-b we have drawn a horizontal dotted line along this value as analysis guidance. In the case of the analysis of the evolution of the binding energy, it is important to obtain both low energies and low standard deviation values, since they indicate a strong and stable binding. From results of peptides of first set in Fig. C2a, we highlight as best binders the peptides 43 (with the lowest average binding energy of the set), 482 (second lowest average energy and lowest standard deviation value), 331 and 79. Otherwise, peptides 50f, 52, 72, 74, 84, 91, 279f, 332, 350, 359, and 460 are discarded since their energies are clearly above

-10 kJ/mol. With respect to the second set of peptides, the best binders according to Fig. C2b are peptides 107 (the lowest average energy of both sets), 58, 122 and 66 (spite of the higher average energy, it has the lowest standard deviation value of both sets indicating the good stability of the binding). Additionally, peptides 76 and 85 are discarded because of the energy values above the limit.



**Figure C2:** (a-b) Evolution of the Binding Energies of the  $\beta 2m$ -peptide complexes of the first and second set. (c-d) Evolution of the distance between the center of masses of the binding site of the  $\beta 2m$  and the peptides of the first and second set.

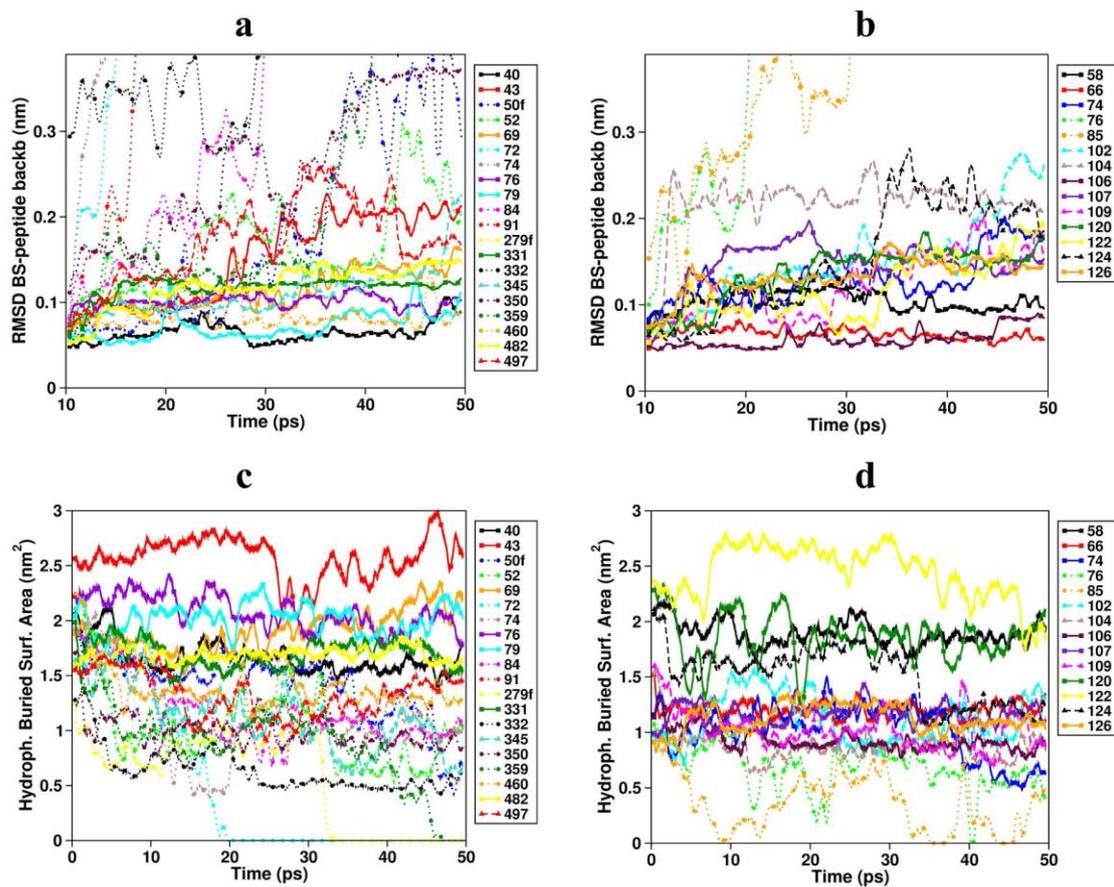
In Fig. C2c-d is shown the evolution of the distance between the BS of  $\beta 2m$  and the peptide for each complex. Since each BS is defined specifically for each complex, it is difficult to compare the average value of the BS-peptide distance between different complexes, although we can establish that average values higher than 1 nm correspond to weak bindings. Nevertheless, the most important criterion in this descriptor is the profile of the curves along the simulation, in which the best is the one with the lowest standard deviation value and without an increasing trend. In the first set in Fig. C2c, curve of peptides

76, 79, 331, 40, 43, 69, and 482 show correct behavior. Moreover, the best peptides in the second set (see Fig. C2d) are 66 (the lowest standard deviation value of all peptides), 58, 74, 107, 106, and 102. In all cases, the standard deviation value is lower than 0.05 nm.

In Fig. C3a-b are plotted the evolution of the RMSD of the BS-peptide backbone complexes of first and second set, respectively. This descriptor show us mainly if there are conformational rearrangements of the  $\beta$ 2m-peptide binding along the simulation when the curve typically reaches a higher plateau. Therefore, the best binders are those without major conformational changes. The curves are started from the geometries at 10 ns to avoid the initial rearrangements in the simulation. From first set in Fig. C3 we chose 40, 76, 79, 331, and 460. From the second set the best binders are peptides 58, 66, and 106.

Additionally, Fig. C3c-d show the evolution of the Hydrophobic Buried Surface Area (HBSA) of peptides of first and second set, respectively. The HBSA is calculated as the difference of the solvent-accessible surface Area (SASA) of the hydrophobic groups of peptide alone and peptide + BS, and evaluates the evolution of the hydrophobic interaction between the  $\beta$ 2m and peptides. In this case, high average values, low standard deviation values and no decreasing trend are the best behavior. From the first set we highlighted peptides 40, 43 (the highest average value of all peptides), 69, 76, 79, 331, and 482. From the second set, we find peptides 58, 66 (the lowest standard deviation value of all peptides), 120, and 122 have the strongest and most stable hydrophobic interaction of the set.

According to these results, we display in the Table C2 the selected sequences of sets, as well as the average values and the standard deviation values of the four descriptors used to perform the screening analysis.



**Figure C3:** (a-b) Evolution of the RMSD between the binding site of the  $\beta 2m$  and the peptides of the first and second set. (c-d) Evolution of the HBSA of peptides of the first and second set.

**Table C2:** Average values and standard deviation values (in parenthesis) of descriptors used for the screening analysis of the selected  $\beta$ 2m-peptide complexes.

Name	Sequence	$E_B$ (kcal/mol)	Dist. (nm)	RMSD (nm)	HBSA (nm <sup>2</sup> )
<b>1st set</b>					
40	CWERQFKLYGKC	-12.79(0.91)	0.75(0.03)	0.06(0.01)	1.66(0.14)
43	CDGYWLRKAQWC	-15.19(1.15)	0.74(0.03)	0.15(0.05)	2.57(0.19)
69	CWERQHQDYTKC	-13.68(0.95)	0.76(0.03)	0.12(0.02)	1.91(0.17)
76	CWERQHQMAYKC	-13.66(1.32)	0.66(0.02)	0.10(0.01)	2.08(0.15)
79	CWERQHQMAYLKC	-13.75(0.73)	0.70(0.02)	0.07(0.02)	2.02(0.13)
331	CFETAWRQNEWC	-13.85(0.93)	0.68(0.04)	0.12(0.01)	1.70(0.12)
482	CYNRVYRKWHRC	-14.87(0.66)	0.76(0.03)	0.12(0.02)	1.69(0.06)
<b>2nd set</b>					
58	CHKMSNIQKNWC	-14.90(0.77)	0.69(0.02)	0.10(0.01)	1.90(0.11)
66	CHDMSNPLKNWC	-11.53(0.55)	0.65(0.01)	0.07(0.01)	1.20(0.06)
74	CHDMSNQLKNWC	-11.73(0.70)	0.69(0.03)	0.13(0.03)	1.02(0.23)
106	CNDALNRLKNWC	-13.43(0.77)	0.73(0.04)	0.06(0.01)	0.93(0.09)
107	CNDALNRLKNQC	-16.23(1.09)	0.64(0.03)	0.14(0.03)	1.16(0.09)
120	CRNMDGNFDQKC	-12.37(1.12)	0.77(0.03)	0.13(0.03)	1.85(0.21)
122	CRNQDGNFDQKC	-14.58(1.06)	0.70(0.04)	0.12(0.04)	2.44(0.24)
126	CNDRLNRLKNQC	-12.55(1.15)	0.69(0.04)	0.13(0.02)	1.13(0.12)

#### C.4 References

1. Kozakov, D., Brenke, R., Comeau, S. R. & Vajda, S. PIPER: An FFT-based protein docking program with pairwise potentials. *Proteins Struct. Funct. Bioinforma.* **65**, 392–406 (2006).
2. Ngan, C.-H. *et al.* FTSite: high accuracy detection of ligand binding sites on unbound protein structures. *Bioinformatics* **28**, 286–287 (2012).
3. Kozakov, D. *et al.* Structural conservation of druggable hot spots in protein-protein interfaces. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 13528–33 (2011).
4. Lavi, A. *et al.* Detection of peptide-binding sites on protein surfaces: The first step toward the modeling and targeting of peptide-mediated interactions. *Proteins Struct. Funct. Bioinforma.* **81**, 2096–2105 (2013).
5. Hess, B., Bekker, H., Berendsen, H. J. & Fraaije, J. G. *Journal of computational chemistry* 18, (12), . **1997**, 1463–1472
6. Trott, O. & Olson, A. J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **31**, 455–61 (2010).
7. Laskowski, R. A. & Swindells, M. B. LigPlot+: Multiple Ligand–Protein Interaction Diagrams for Drug Discovery. *J. Chem. Inf. Model.* **51**, 2778–2786 (2011).

## List of figures

Figure 1.0	
Depicts the double structure of B-DNA with vital segments like minor and major groove that occurs where the back bones are closer and far apart respectively .....	2
Figure 1.1	
Types of immobilization of DNA molecules on solid surfaces that are regularly used in DNA nanotechnology field.....	4
Figure 1.2	
Chemical photolithography for the construction of spatially confined DNA compartments.....	7
Figure 1.3	
Dip pen nanolithography .....	8
Figure 1.4	
Construction of confined DNA compartment via Nanografting .....	10
Figure 1.5	
The effect of probe density on hybridization.....	13
Figure 1.6	
Fine-tuning the parking density of confined DNA nanostructures during nanografting process .....	14
Figure 1.7	
Mechanical responses of highly dense confined DNA nanostructures before and after <i>in situ</i> hybridization process .....	16
Figure 1.8	
Effect of ionic strength on the conformation and height of spatially confined dsDNA brushes .....	19
Figure 1.9	
Detection of two-dimension enzyme diffusion through confined dsDNA nanopatches .....	21
Figure 1.10	
Real time digestion of endonucleases within density-gradient DNA brush .....	23
Figure 1.11	
Combination of DDI and nanografting.....	25

Figure 2.0	
	Schematic representation of the action of type I and type II endonuclease enzymes..43
Figure 2.1	
	Generation of three distinct sequence-dependent, spatially resolve DNA nanoreactors, via nanografting .....46
Figure 2.2	
	Construction of highly dense DNA nanoreactors via spontaneous formation of self-assembled monolayer of pre-hybridized dsDNA molecules on ultra-flat gold substrate .....48
Figure 2.3	
	Sequence-density related function of BamHI revealed by AFM height measurements ..... 50
Figure 2.4	
	Sequence-density related activity of Bam HI within self-assembled monolayer of dsDNA made up of DNA-3 .....53
Figure 2.5	
	Side-side AFM height profiling of DNA nanoreactors made up of DNA1 (left) and DNA 3(right) .....55
Figure 2.6	
	Examining the action of BamHI within DNA 1 (left) and DNA 3(right) molecules at upper intermediate densities..... 57
Figure 2.7	
	Correlation plots of final relative height versus initial relative height ..... 59
Figure 2.8	
	Correlation plots of relative final height versus relative initial height, showing the unexpected action of BamHI within DNA 3 monolayers, revealed by side-by-side AFM height measurements ..... 60
Figure 3.0	
	(a) Structure of $\beta$ 2-microglobulin (gray) bound to the optimized peptide 461 (blue). The $\beta$ 2m binding site is composed by A, B, E, D $\beta$ -sheet-forming strands (red). The $\beta$ 2m lysine residues are also highlighted (yellow). (b) Evolution of the scoring energies of the three Monte Carlo replicas ( $T=0.3, 0.6, 0.9$ ) during a typical run. Low score and soluble peptides (green crosses) are selected for successive screening. Selected configurations along the optimization path are shown in the insets ..... 70

Figure 3.1	
	Primary screening: MD simulation results. (a) Average scoring energy, (b) Separation distance between $\beta$ 2m binding site (BS- $\beta$ 2m) and peptide, and (c) RMSD of BS- $\beta$ 2m+peptide along the simulation time ..... 71
Figure 3.2	
	AFM-based approach for immobilization of cyclic peptide and detection of $\beta$ 2m-cyclic peptide interactions on surfaces..... 75
Figure 3.3	
	Shows an array of nanoassemblages, each generated with different surface coverage (S/A), from the least (left column) to the highest (far right) ..... 76
Figure 3.4	
	AFM height measurements as a function of surface coverage ..... 77
Figure 3.5	
	AFM-based quantitative approach for analysing surface-based DNA-pep461 assemblages before and after $\beta$ 2m recognition at different target ( $\beta$ 2m) concentration ..... 80
Figure 3.6	
	An example of a subsequent height measurement procedure employed in determining the average height difference across the topographic AFM micrograph of DNA-pep461 assemblage before and after $\beta$ 2m recognition (100 $\mu$ M concentration)..... 81
Figure 3.7	
	Graph of overall differential height associated with $\beta$ 2m-cyclic peptide interactions as a function of $\beta$ 2m concentration. The experimental data (closed black spheres) was fitted using Hill's equation ..... 82
Figure 4.1	
	Peptimap graphical output representing a selection of the identified binding sites. These are all located in the regions (a) A, (b) B, (c) C, (d) D, (e) E and (f) F of $\beta$ 2m and the representation on the far right shows the schematic of the different labeled regions of $\beta$ 2m ..... 87
Figure 4.2	
	Simulation snapshot of (a) the pep331 and (b) pep381 with $\beta$ 2m, each binding a different solvent exposed binding site on $\beta$ 2m..... 88

Figure 4.3	
	Using AFM topographic height measurement to detect the recognition of DNA-peptide assemblages (DNA-pep331) towards $\beta$ 2m at 38 $\mu$ M concentrations..... 90
Figure 4.4	
	Recognition of $\beta$ 2m by DNA-pep381 assemblages at the 7 $\mu$ M concentration of target solution. The recognition was detected via label-free AFM topographic height measurements ..... 91
Figure 4.5	
	Shows the serial height measurement procedure with one-to-one profile mapping employed in determining the average height difference across the topographic AFM micrographs of DNA-pep331 assemblage before and after $\beta$ 2m recognition (7 $\mu$ M concentrations). Scale bar, 750 nm..... 92
Figure 4.6	
	Subsequent height measurement procedure employed in determining the average height difference across the topographic AFM micrographs of DNA-pep381 assemblage before and after $\beta$ 2m recognition (at 7 $\mu$ M concentration). The scan size of the AFM micrographs is 3 x 3 $\mu$ m <sup>2</sup> ..... 93
Figure A.1	
	pH-controlled self-assembly of DNA tiles ..... 106
Figure A.2	
	(a) Upstream pH-dependent DNA circuit. (b) Triplex formation in the pH-dependent substrate complex was studied by incorporating a pH-insensitive FRET pair at the ends of the clamp-like triplex forming strand (left), and measuring the fluorescence signal at different pHs (right). Stable triplex formation is observed only at pHs below 7.0. (c) The pH-dependent triplex complex in the substrate inhibits strand displacement reaction upon catalyst addition (pH 5.0, left) ..... 107
Figure A.3	
	pH-dependent self-assembly of DNA tiles ..... 109
Figure A.4	
	Kinetic of pH-dependent self-assembly of DNA tiles ..... 110

Figure A.5	
pH independent DNA nanotubes weaved into DNA-mat revealed by the AFM characterization .....	112
Figure SI1	
Reactive and protected tile .....	116
Figure SI2	
Upstream pH-dependent strand displacement circuit characterization.....	116
Figure SI3	
Upstream pH-dependent strand displacement circuit characterization.....	117
Figure SI4	
Characterization of the catalytic efficiency of the upstream circuit upon the addition of the Catalyst.....	118
Figure SI5	
pH-independent strand displacement circuit (Control).....	119
Figure SI6	
pH-independent strand displacement circuit (Control).....	120
Figure SI7	
pH-independent strand displacement circuit (Control).....	121
Figure SI8	
Statistical analysis of assembled tiles length .....	122
Figure SI9	
Control of pH-dependence of DNA tile self-assembly reaction .....	123
Figure SI10	
Fluorescence microscopy images of the integrated reaction network in the absence of Catalyst.....	124
Figure SI11	
Kinetic of the pH-dependent nanotubes self-assembly.....	125
Figure S112	
AFM characterization after 4 days of incubation.....	125
Figure B1	
(a) Secondary Screening: SPR signals of eleven selected peptides at six different concentrations. (b) SPR sensogram of the selected sequence (sequence 461, in the inset). .....	129

βFigure B2

SPR sensogram for cyclic peptides that undergo secondary screening using SPR ... 129

Figure B3

ITC Calorimetric (a) raw and (b) integrated data for β2m (42 μM) titration with peptide 461 (1 mM). Data fitting was achieved by ‘one set of sites’ model..... 131

Figure B4

Representative chemical shift deviations measured from 2D [<sup>1</sup>H, <sup>15</sup>N] HSQC experiments. Lines are only intended to guide the eye..... 132

Figure B5

a) Overlay of 2D [<sup>1</sup>H, <sup>15</sup>N] Heteronuclear Single-Quantum Correlation (HSQC) spectra of <sup>15</sup>N-labeled β2m 45 μM in the free form in blue and in the presence of peptide 461 in red (peptide/protein = 36.5/1). The peaks presenting the largest chemical shift deviations are boxed in red and enlarged in the excerpts. The signals affected by the most consistent intensity variations are highlighted with red circles. (b) NMR results depicted on β2m structure: residues showing larger chemical shift deviations and intensity variations in the 2D [<sup>1</sup>H, <sup>15</sup>N] HSQC experiments are colored in red and in blue, respectively. (c) Nitrogen root mean square fluctuation difference (ΔN-RMSF) between free and peptide bound β2m with vertical lines indicating amino acids presenting either chemical shift (red) or intensity increase (blue) upon binding, averages over 250 ns MD trajectories. (d) Peptide root mean square deviation (RMSD) at fixed β2m binding site (blue) and distance between the peptide and its binding site (black) along the MD trajectory (left panel) and their distributions (right panel)..... 133

Figure B6

Primary screening: MD simulation results. Selected peptides are highlighted by thicker lines. (a) Average scoring energy, (b) Separation distance between β2m binding site (BS-β2m) and peptide, and (c) RMSD of BS- β2m+peptide along the simulation time..... 135

Figure C1

Evolution of RMSD of β2m taking as a reference (a) the 0 ns and (b) 12.5 ns structures. .... 139

Figure C2

(a-b) Evolution of the Binding Energies of the  $\beta$ 2m-peptide complexes of the first and second set. (c-d) Evolution of the distance between the center of masses of the binding site of the  $\beta$ 2m and the peptides of the first and second set ..... 142

Figure C3

(a-b) Evolution of the RMSD between the binding site of the  $\beta$ 2m and the peptides of the first and second set. (c-d) Evolution of the HBSA of peptides of the first and second set..... 144

## List of tables

### Table 2.0

DNA sequences for DNA-1, DNA-2, and DNA-3 ..... 62

### Table 4.0

Ranking of different binding sites obtained for each  $\beta$ 2m conformation..... 87

### Table B1

Peptide sequences selected through SPR secondary screening, pI (isoelectric point) and dissociation constants toward  $\beta$ 2m are also reported ( $K_d$ ). Highlighted sequences (italics, red), tagged 461 and 381, are the peptides used for experimental validation by AFM in chapter 3 and 4 respectively..... 130

### Table B2

Scoring Energies (E, in kcal/mol) and sequences of the soluble peptides with the lowest Scoring E. The peptides are the output of two runs of the algorithm. Each run consists of three parallel replica exchange Monte Carlo optimisation run at the pseudotemperature  $T=0.3, 0.6, 0.9$ . The solubility has been predicted with Innovagen's peptide calculator webtool [PepCalc.com. <http://pepcalc.com/>. Innovagen AB, 2012.]. Note the sequences used for experimental validation via AFM are highlighted in italics, red. .... 134

### Table C1

Solvent Accessible Surface Area (SASA) of each most probable binding site and sequences of residues involved in the most probable binding sites of every region 140.

### Table C2

Average values and standard deviation values (in parenthesis) of descriptors used for the screening analysis of the selected  $\beta$ 2m-peptide complexes..... 145