UNIVERSITY OF SOUTHAMPTON

FACULTY OF ENGINEERING AND PHYSICAL SCIENCES

School of Chemistry

Determining Surface Oriented SRCD Signatures of Proteins on DNA Nanostructures

by

Christina Xyrafaki

Thesis for the degree of Doctor of Philosophy

June 2021

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF ENGINEERING AND PHYSICAL SCIENCES
School of Chemistry

Doctor of Philosophy

DETERMINING SURFACE ORIENTED SRCD SIGNATURES OF PROTEINS ON DNA NANOSTRUCTURES

by Christina Xyrafaki

Enzyme and protein activity is affected by their relative orientation when immobilised on surfaces. Localising proteins on a well-behaved surface makes it possible to study their orientation and by extension, their structure, which is pursued using oriented Synchrotron Radiation Circular Dichroism (SRCD) spectroscopy.

SRCD is a powerful and sensitive tool to investigate the structure of proteins. Using the B23 beamline at Diamond Light Source and the novel CD imaging (CDi), methods for the preparation and measurement of oriented SRCD samples are developed in order to characterise them on surfaces. Proteins bound to two-dimensional (2D) DNA nanostructures are analysed in solution as well as absorbed on glass surfaces as films. The 2D DNA nanostructures are designed to form long-range arrays for consistent orientation over the entire sample area. Both the DNA on its own and functionalised with protein are studied in order to investigate if the alignment of the proteins on the DNA lattices can result in the change of the SRCD signature obtained from the proteins.

For this thesis, two DNA two-tiled systems are chosen, which self-assemble into 2D nanogrids and double-crossover (DX) arrays. As proof-of-concept proteins, Green Fluorescent Protein (GFP) and Cytochrome b_{562} are used which are site-specifically modified with azidophenylalanine, which allows to precisely position an azide at a defined residue for a single site DNA attachment using click chemistry.

Finally, an new approach of sample preparation is investigated as an alternative to the film formation, which is based on the immobilisation of biomolecules in a reversible manner and which relies on non-covalent fluorous-fluorous interactions. This method can overcome challenges that are faced during film formation which is particularly important for protein studies under dry conditions.

Contents

Abbreviations **xxi**

Nomenclature **xxiii**

Declaration of Authorship **xxv**

Acknowledgements **xxvii**

1 Introduction **1**
1.1 Proteins . . . 1
1.1.1 Protein structure . . . 1
1.1.2 Current studies of proteins on surfaces . . . 4
1.2 Deoxyribonucleic Acid, DNA . . . 6
1.2.1 Building blocks of nucleic acids and tertiary structure of DNA . . . 6
1.2.2 Oligonucleotide Synthesis . . . 8
1.3 DNA as structural material . . . 10
1.3.1 DNA nanostructures as templates for self-assembly . . . 11
1.4 Protein-DNA conjugation . . . 13
1.5 Introduction to circular dichroism and synchrotron radiation circular dichroism . . . 15
1.5.1 Circular Dichroism (CD) and applications . . . 15
1.5.2 Synchrotron Radiation Circular Dichroism (SRCD) Spectroscopy . . . 19
1.6 Atomic Force Microscopy, AFM . . . 20
1.7 Aims and objectives . . . 22
1.8 Thesis structure . . . 22

2 2D DNA Nanogrid for Protein Attachment **25**
2.1 2D DNA Nanogrid Design Overview . . . 26
2.2 Results and discussion . . . 26
2.2.1 Synthesis and analysis of the oligonucleotides . . . 26
2.2.2 Formation of 2D DNA cross-tiles and nanogrid . . . 29
2.2.3 Synthesis of the modified oligonucleotide . . . 31
2.2.4 Protein Modification- Click Chemistry . . . 33
2.2.5 Protein-ssDNA co-assembly on DNA nanogrid . . . 36
2.2.6 Atomic Force Microscopy of 2D DNA nanogrid . . . 37
2.2.6.1 AFM under dry conditions . . . 38
2.2.6.2 Liquid AFM . . . 39
2.2.7 SRCD & CD Spectroscopy for the 2D DNA crosstiles & nanogrid . . . 40

2.2.7.1 SRCD measurements in solution . . . 41
2.2.7.2 Benchtop CD measurements in solid phase . . . 43
2.2.7.3 CD *imaging* . . . 46
2.3 Discussion and conclusion . . . 56

3 DNA Double-Crossover array for protein attachment 59
3.1 DNA Double-Crossover (DX) Array Design Overview . . . 60
3.2 Results and discussion . . . 63
3.2.1 Synthesis and analysis of the oligonucleotides . . . 63
3.2.2 Formation of the DX tiles and array . . . 66
3.2.3 Protein Modification- Click Chemistry . . . 67
3.2.4 Protein-ssDNA attachement to DX tile and array . . . 70
3.2.5 Atomic Force Microscopy of DX-Array with protein attachment . . 72
3.2.5.1 Liquid AFM of DX-Array with GFP attachment . . . 72
3.2.5.2 AFM under dry conditions of DX-Array with Cytb_{562} attachment . . . 73
3.2.6 CD & SRCD Spectroscopy of the DX tiles and DX-Array . . . 74
3.2.6.1 Thermal denaturation of cytb_{562} using circular dichroism 74
3.2.6.2 CD measurements in solution . . . 77
3.2.6.3 CD experiments using films/ CD *imaging* . . . 86
3.3 Discussion and conclusion . . . 98

4 DX-array immobilisation on fused silica surface using fluorous effect 101
4.1 Fluorous Chemistry . . . 102
4.2 Results and discussion . . . 103
4.2.1 Fabrication of fused silica surface using fluorous tags . . . 103
4.2.2 Formation of fluorous-tagged DX-array . . . 105
4.2.3 Immobilisation of the fluorous-tagged DX-array on the fluorous modified fused silica surface . . . 109
4.2.4 CD and SRCD measurements of the fluorous-tagged DX-array on fused silica surfaces . . . 111
4.3 Discussion and conclusion . . . 123

5 Conclusions & Future Work 125

6 General Experimental Procedures 129
6.1 Materials & Instruments . . . 129
6.2 DNA Synthesis . . . 130
6.3 Glen Pak Cartridge Purification . . . 130
6.4 High Performance Liquid Chromatography . . . 131
6.5 Denaturing Urea Gel-PAGE (DNA Purification) . . . 132
6.6 DNA modification with 5' BCN-CEP II modifier . . . 133
6.7 Desalting Procedure of BCN-DNA (Glen-Pak) . . . 133
6.8 Protein-DNA modification (Click Chemistry) . . . 133
6.9 SDS-PAGE . . . 133
6.10 DNA nanostructures' annealing protocols in solution . . . 135
6.11 Agarose Gel 1% . . . 137
6.12 Protein-ssDNA Co-Assembly on DNA nanostrustures . . . 137

6.13 Sample preparation for AFM 138
6.14 CD and SRCD Spectroscopy 138
6.15 Immobilisation of DNA nanostructures on fluorous modified quartz disk . 139
6.16 Removal of the DNA nanostructure from the fluorous modified quartz surface and its re-immobilisation 139

References **141**

Appendices **155**

Appendix A **157**

Appendix B **177**

Appendix C **187**

List of Figures

1.1 Central dogma of molecular biology- protein synthesis. Reprinted from[1] . 1
1.2 Levels of structure in proteins. Reprinted from[6] 3
1.3 Schematic diagram showing the attachment of the protein on the Au nanoparticle. Tetracysteine motif. Reprinted from[18] 5
1.4 Schematic diagram of IgG immobilisation on a gold surface modified with solution of CS_2 and protein A. A gold slide is firstly immersed into the mixed solution of CS_2 and protein A and then, incubated in IgG solution to form the IgG sensing surface. In the final step, the incubation of the sensing surface in anti-IgG solution will allow the evaluation of the bioactivity of the immobilised IgG. Reprinted from[19] 5
1.5 **a)** The building blocks of nucleic acids: phosphate group (green), sugar unit (red), nucleobases(blue). **b)** Structure of four nucleobases. 7
1.6 Watson-Crick Base pairings A-T and C-G. dR indicates the position of the deoxyribose. 7
1.7 Three main forms of DNA, from left to right: A-DNA, B-DNA, Z-DNA. Reprinted from[31] 8
1.8 Different helical parameters for the common DNA conformations. Reprinted from[33] 8
1.9 The phosphoramidite oligonucleotide synthesis cycle (ATDBio Nucleic Acid book). Reprinted from[39] 9
1.10 Four arm-DNA branch junctions with complementary sticky ends (a-a' and b-b' pairing) to self-assemble into 2D lattice. Reprinted from[42] 10
1.11 From left to right: double-crossover (DX) tiles, triple-crossover (TX) tiles, 4x4 tile and three-point-star tile. The top images are the structures of motifs, the middle images 2D and 3D assembly patterns, and the bottom images are AFM images of the lattices. Adapted from[46] 11
1.12 DNA origami structure. **(A)** Method employed to prepare a DNA origami structure. **(B)** Design of a self-assembled DNA origami structure and geometry of the incorporated dsDNAs. **(C)** Design and AFM images of self-assembled DNA origami structures. Adapted from[49] 12
1.13 **(A)** Biotin-monoavidin binding. **(B)** C-terminal fused intein protein-DNA conjugation by expressed protein ligation. **(C)** Acyl transfer reaction catalysed by microbial tansglutaminase (MTG) between N-carbo benzyloxyglutaminylglycine (Z-QG) DNA and short peptide tag. The schemes were adapted from[46] 13

1.14 Schematic of the genetically encoded β-lactamase with the azide in different residues and its further modification with the ssDNA through biocompatible and orthogonal click chemistry. Finally its attachment to the DNA origami tiles. The conjugation position and by extension the orientation of β-lactamase to the DNA origami tile is important in terms of catalytic activity. In the case at the top, this conjugation position leads to greatly improved enzymatic activity than the free enzyme and in the case at the bottom, it shows lower activity and therefore, product transformation. Reprinted from[24] 14
1.15 CD signatures of α-helical protein (blue), β-strand (red) and unordered (black) proteins. Reprinted from[7] 16
1.16 Circular dichroism spectra for right-handed A-RNA and A-DNA, right-handed B-DNA, and left-handed Z-DNA. Adapted from[85] 18
1.17 Optical characterisation of the DNA origami dressed with single AuNRs. Chiral molecule-induced CD is clearly observable centred at around 730 nm. CD in the parallel AuNR cases (red and blue) shows opposite chirality compared to that in the vertical AuNR case (purple). The experimental results clearly demonstrate that the chiral molecule induced CD is highly orientation dependent. Reprinted from[92] 19
1.18 **a)**, Unique B23 vertical sample chamber equipped with a motorised XY stage for imaging sample films. **b)**, CD *imaging* (CDi) of a polymer film using B23. Reprinted from[99] 20
1.19 AFM schematic. Reprinted from[102] 21

2.1 **a)** Two tile system to form **b)** the 2D nanogrid. In Subfigure (a), the single strand in red is the extended strand for the conjugation of proteins and enzyme. # 1-8 and # 1'-8' are the designated sticky ends. Reprinted from[52] 27
2.2 **Left**, Schematic of cross tile A (CTA), which consists of 10 strands. **Right**, Schematic of cross tile B (CTB), which consists of 9 strands. Reprinted from.[52] 27
2.3 UV-Vis spectra of unmodified DNA with a peak at 260 nm for DNA. UV-Vis spectra of the strands involved in CTA tile **(left)** and in CTB tile **(right)**. 29
2.4 1.0% agarose gel. Lane 1: AB grid annealed with Protocol F1, Lane 2: AB grid annealed with Protocol F2, Lane 3: CTA Tile, Lane 4: CTA Tile, Lane 5: CTB, Lane 6: Control a consists of 4 strands involved in tile CTA, but not heated. 31
2.5 Strained bicyclononyne (BCN). 31
2.6 The ssDNA with the terminal bicyclononyne (BCN) for subsequent copper-free click elaboration. 32
2.7 UV-Vis spectra of BCN - DNA with a peak at 265 nm. 33
2.8 Structure of sfGFP^{azF204}. Residues critical to function are shown as spheres, and residue targeted for replacement with azF is shown in stick representation. Adapted from[110] 34
2.9 The protein contains a genetically encoded p-azido-L-phenylalanine (azF; enclosed in box) and the ssDNA contains the BCN. The two components create a triazole link. 34

2.10 Gel mobility shift analysis of protein-DNA conjugation. The left hand lane is the molecular weight standards. Lanes represented the unmodified protein and the protein-DNA are annotated on the figure. Three different concentrations were used for the protein-DNA. The MW of the sfGFP^{azF204} and the sfGFP^{azF204}-DNA is ∼27.7 kDa and ∼37 kDa, respectively. 35
2.11 **Left,** UV-Vis spectrum of the protein-DNA with a peak at 265 nm for the DNA and at 495 nm for the sfGFP^{azF204} and UV-Vis spectrum of BCN-DNA with a peak at 265 nm, **Right,** Fluorescence of protein DNA, emission spectrum was recorded on excitation at 485 nm. 35
2.12 **(a)** The two tile system to form the 2D DNA nanogrids. **(b)** The 2D DNA nanogrid with the extended strand on each A tile pointing out of the plane. **(c)** Assembly of the protein on the DNA grids. Adapted from[52] 36
2.13 Different modes of AFM, including contact and tapping mode. Adapted from[113] 37
2.14 **a),b)** AFM images of 2D DNA nanogrid (1 μM)as captured using AFM in air (1 x 1 μm). **c)** AFM image of 2D DNA nanogrid as captured by Yan *et al.* (190 x 190 nm).[52] 38
2.15 Images of 1 μM DNA nanogrid with protein attached **A.** Image 1.3 x 1.3 μm. Cross structures are visible and eclosed in blue squares. **B.** Image 2.0 x 2.0 μm. The protein dots are more visible here than the nanogrid. The section analysis presented is for the zoom-in area that shows 3 protein dots with a distance of 38 nm from each other. 39
2.16 Images of 1 μM 2D DNA nanogrid as captured in liquid AFM. **a)** Image 791 x 791 nm. **b)** Image 1.4 x 1.4 μm. **c)** Image 1.0 x 1.0 μm. 40
2.17 Image of 1 μM 2D DNA nanogrid (1.4 x 1.4 μm) from Figure2.16(b) as captured in liquid AFM and the analysis sections for each zoomed area **A.** A zoom-in area showing a square lattice. The section analysis above represents the diagonal distance of the tiles along the yellow arrow. **B.** Another zoom-in area showing a bigger square lattice. The analysis sections below and above describes the linear distance of the tiles along the red and blue arrows, respectively. **C.** The expected distances of the cross tiles (blue-to-blue tile, diagonal repeat: 25-27 nm and linear repeat: 35-39 nm). DNA nanogrid scheme adapted from ref.[52] 41
2.18 SRCD spectra of the CTA (red) and CTB (blue) tiles and the AB grids Protocol F1 & F2 (green and blue, respectively) in solution. Conditions: 1 μM in 1x TAE/Mg^{2+} buffer. 42
2.19 High tension (HT) on the photomultiplier detectors of Module B of the CTA (red) and CTB (blue) tiles and the AB grids Protocol F1 & F2 (green and blue, respectively) in solution. 43
2.20 **Left**, CD spectra of AB grid films before and after 90° rotation (solid and dotted lines, respectively). **Right**, Absorbance spectra of AB grid films before and after 90° rotation (solid and dotted lines, respectively). 44
2.21 CD spectra of AB grid films before and after 90° rotation (solid and dotted lines, respectively) after offset at around 310 to 320 nm. 44
2.22 **Left**, CD spectra of AB grid film (Protocol F1), upon 8 rotations by 45° clockwise. **Right**, Absorbance spectra of AB grid film, upon 8 rotations by 45° clockwise. 46

2.23 Images as captured using ZEISS Stereo Microscope. **a)** Film of the AB grid (Protocol F1). **b)** Film of the AB grid (Protocol F2). 46
2.24 Unique B23 vertical sample chamber equipped with a motorised XY stage (Prior Stage) for imaging sample films. 47
2.25 **Left**, SRCD data of the film for the CTB tile, mapping a 4 x 4 grid array of 700 μm^2 step size. **Right**, Absorbance data of the film for the B tile, mapping a 4 x 4 grid array of 700 μm^2 step size. 48
2.26 Multi-panel showing on the **left,** SRCD data of the film for the CTB tile and on the **right** absorbance data of the film for the B tile, after mapping a 4 x 4 grid array of 700 μm^2 step size. 49
2.27 **a)** SRCD data of AB film (Protocol F2) after scanning 3x3 grid array at 500 μm interval steps including also an image of the film as captured with the chamber's camera, showing the irradiated 3x3 sampling area. The 2D maps were generated for the 283 and 245 nm, respectively. **b)** SRCD data and 2D maps (283 and 245 nm) upon rotation of the sample specimen by 90° rotation clockwise. 50
2.28 **a)** SRCD data of AB film (Protocol F1) after CD*i* experiment by scanning 4x4 grid array of 700 μm. The 2D colour maps for the fixed wavelengths 280 and 240 nm are also presented. **b)** Absorbance data after CD*i* experiment by scanning 4x4 grid array of 700 μm and a 2D colour map at 263 nm. 52
2.29 Multi-panel (square grid map) showing on the **left,** SRCD data of the film for the AB grid (Protocol F1) and on the **right** absorbance data of the film for the AB grid (Protocol F1), after mapping a 4 x 4 grid array of 700 μm^2 step size. 53
2.30 G-factor spectra of AB grid (Protocol F1) in film. 54
2.31 SRCD data of the film for the AB grid with GFP attached after mapping of a 4x4 grid array of 700 μm step size. The two SRCD graphs on the right are zoomed in a scale of 190-350 nm and 300-600nm, respectively. . 55
2.32 UV absorbance data of the film for the AB grid with GFP attached after mapping a 4x4 grid array of 700 μm step size. The graph on the right side is the zoomed graph data at the DNA wavelength range (characteristic peak of DNA at 260 nm). 56

3.1 **a)** Model structure of DAE class of DX molecules. Adapted from[103] **b)** DX tiles and arrays. The two tiles can be mixed in equal amounts to generate the DX-AB array. Adapted from[122] 60
3.2 The DX tiles (grey and blue) that generate the DX-array. Their sticky ends are presented as oddly shaped adaptors which interlock to specific pieces of the other tile in order to form the bigger DX molecule 61
3.3 **a)** Original DX-A tile with sequences included. Adapted from[122] **b)** and **c)** Modified DX-A tiles, DX-$A_{5'}$ and DX-$A_{3'}$ with the extended ssDNA (red) either on the 5' prime or the 3' prime end, respectively. 62
3.4 DX tiles and the DX-array. The DX-A tile (grey) and DX-B tile (blue) contain complementary sticky ends and then can be mixed in equal amounts to generate the DX-array. The DX-A consist of the extended strand (red), the anchoring point for the protein attachment. 63
3.5 DX-B tile with sequences included. 63

3.6 UV-Vis spectra of unmodified DNA with a peak at 260 nm for DNA. UV-Vis spectra of the strands involved in DX-A tile **(left)** and in DX-B tile **(right)**. 65
3.7 1.0 % agarose gel. Lane 1: Control a1, Lane 2: DX-$A_{5'}$, Lane 3: Control a, Lane 4: DX-$Array_{5'}$, Lane 5: Control b1, Lane 6: DX-B, Lane 7: Control b. 66
3.8 **Left,** short axis (SA) cytochrome b_{562} with azF mutation at position 5 and cysteine mutation at position 104. **Right,** long axis (LA) cytochrome b_{562} with azF position at 50 and cysteine mutation at position 21. 68
3.9 Gel mobility shift analysis of protein-DNA conjugation. The left hand lane is the molecular weight standards. Lanes represented the unmodified variants of cytb_{562} and the protein-DNA are annotated on the figure. The monomeric and the dimeric form of the variants are visible. This SDS-PAGE image was obtained by our collaborator Benjamin Bowen. 68
3.10 **Left**, UV-Vis spectra of SA cytb_{562} dimer, SA cytb_{562} monomer and SA cytb_{562}+ssDNA with a peak at 265 nm for the DNA and at 418 nm for the cytb_{562}. **Right**, UV-Vis spectra of LA cytb_{562} dimer, LA cytb_{562} monomer and LA cytb_{562}+ssDNA with a peak at 265 nm for the DNA and at 418 nm for the cytb_{562}. The UV-Vis data were obtained by our collaborator Benjamin Bowen. 69
3.11 **a)** Short axis cytochrome b_{562} (Q71) including the location of the azF mutation on the sort axis of the protein at position 71. **b)**, Gel mobility shift analysis of protein-DNA conjugation. The left hand lane is the molecular weight standards. Lanes from left to right show bands of the unmodified and the modified variant of Q71 cytb_{562}. The SDS-PAGE image was obtained by our collaborator Benjamin Bowen. 69
3.12 Schematic representation of the final DX-array system. DX-A tile and DX-B tile are mixed in equal amounts to generate the DX-array, which includes the extended strand on each DX-A tile (either on the 5' or 3' end of the middle DNA strand) pointing out of the plane. Finally, the protein of interest is immobilised on the DX-array either close to the array surface or on top of it. 70
3.13 **a)** 1.0% agarose gel of the DX-$A_{5'}$ tile and DX-$Array_{5'}$ with and without protein (SA cytb_{562}, azF in position 5). Ration of protein:DNA=1:1. **b)** 1.0% agarose gel of the DX-$A_{3' \text{ or } 5'}$ tile and DX-$Array_{3' \text{ or } 5'}$ with and without protein (Q71 cytb_{562}). Ration of protein:DNA=2:1. 71
3.14 **a)** AFM image of DX-array with GFP attached (1:0.7 ratio) as captured wih liquid AFM, 1.2 μm scan size.**b)** DAE-O AB lattice with 33±3 nm periodicity, 500nm scan size, captured by Seeman *et al.*[103] 72
3.15 **a)** AFM image of 1 μM DX-array with cytb_{562} attached and the section analysis of the light blue dotted arrow showing the distance between proteins, scan area 302.6 nm **b)** AFM image of sample containing the DX-array, the bioti-TFO and an excess of SA as captured by Rusling, 300 nm scan size. Adapted from[122] **c)** The expected distances between tile to tile. 73
3.16 **a)** CD data of the LA cytb_{562}-ssDNA, full wavelength. **b)** CD data of the LA cytb_{562}-ssDNA, heme region and, **c)** absorbance data as they were all recorded during thermal denaturation experiment. 75

3.17 Melting temperature of the LA cytb_{562}-ssDNA at 55.5 °C as calculated using Boltzmann fit analysis in Origin. 76
3.18 Conventional CD **(a, b)** and UV absorbance data **(c)** of the DX-A tile, DX-array and of the different variants of cytb_{562} before and after their hybridisation on the tile or array as captured in solution. 78
3.19 Conventional CD **(a)** and UV absorbance **(b)** data of the Q71 cytb_{562} before and after hybridisation to the DNA nanostructure. The protein/DNA region (200-350 nm) was scanned with 3.0 mm path length cuvette. 80
3.20 Conventional CD **(a)** and UV absorbance **(b)** data of the Q71 cytb_{562} before and after hybridisation to the DNA nanostructure. The heme region (350-500 nm) was scanned with 1.0 cm length cuvette. 81
3.21 CD data of the Q71 cytb_{562} after the DX-A$_{3'\text{ or }5'}$ or DX-array$_{3'\text{ or }5'}$ subtraction as recorded at the protein/DNA region (a) and heme region (b). The solid fainted lines belong to the CD signal of the Q71 cytb_{562}azF (pink) and Q71 cytb_{562}-ssDNA (purple) used for comparison. 82
3.22 Conventional CD **(a)** and UV absorbance **(b)** data of the Q50 cytb_{562} before and after hybridisation to the DNA nanostructure. The protein/DNA region (200-350 nm) was scanned with 3.0 mm path length cuvette. 84
3.23 Conventional CD **(a)** and UV absorbance **(b)** data of the Q50 cytb_{562} before and after hybridisation to the DNA nanostructure. The heme region (350-500 nm) was scanned with 1.0 cm length cuvette. 85
3.24 CD data of the Q50 cytb_{562} after the DX-A$_{3'\text{ or }5'}$ or DX-array$_{3'\text{ or }5'}$ subtraction as recorded at the protein/DNA region (a) and heme region (b). The solid fainted lines belong to the CD signal of the Q50 cytb_{562}azF (orange) and Q50 cytb_{562}-ssDNA (green) used for comparison. 86
3.25 SRCD (left) and UV absorbance (right)) data of the DX-A tile (1 μM). Two scans (scan 1, solid and scan 2 dashed) were recorded for each spot of the film 87
3.26 **a)** SRCD spectra of DX-array film after scanning a 4x4 grid array of 500 μm step size. An image of the film as captured with the camera of Linkam stage, showing the irradiated 4x4 sampling area. **b)** 2D and 3D colour maps at 220 and 275 nm (left and right, respectively). 89
3.27 Square grid maps of SRCD spectra (left) and UV-absorbance spectra obtained after mapping a 4x4 grid array area with 500 μm step size. . . . 90
3.28 **a)** UV-absorbance spectra of DX-array film after scanning a 4x4 grid array of 500 μm step size. An image of the film as captured with the camera of Linkam stage, showing the irradiated 4x4 sampling area. **b)** 2D and 3D colour maps at 220 and 275 nm (left and right, respectively). . 91
3.29 SRCD data. Image plots and 3D colour maps after single wavelength mapping of 50 x 50 points with 50 μm per step at **a)** 220nm and **b)** 275 nm. 92
3.30 G-factor spectra for the DX-array film. 93
3.31 SRCD spectra showing the average of three scans for **a)** SA cytb_{562}-ssDNA at 4 (left) and 10 μM (right) and **b)** LA cytb_{562}-ssDNA at 7 (left) and 10 μM (right). 94
3.32 SRCD spectra for the films of the different Q71 cytb_{562} samples.**a)** Cytb_{562}-azF, **b)** Cytb_{562}-ssDNA, **c)** DX-A$_{(3')}$+cytb_{562} and **d)** DX-A$_{(5')}$+cytb_{562}. All samples are in 1 μM concentration and various positions of each film were probed. 95

3.33 UV absorbance spectra for the films of the different Q71 cytb_{562} samples.**a)** Cytb_{562}-azF, **b)** Cytb_{562}-ssDNA, **c)** DX-A$_{(3')}$+cytb_{562} and **d)** DX-A$_{(5')}$+cytb_{562}. 96
3.34 SRCD and UV absorbance data of **a)** DX-array$_{(3')}$+cytb_{562} film after scanning a 3x3 grid array of 500 μm step size and **b)** DX-array$_{(5')}$+cytb_{562} film after scanning a 10x10 grid array of 150 μm step size. 97
3.35 SCRD spectra of DX-array$_{(3')}$+cytb_{562} as the sample goes from **a)** liquid form to **b)** half dry/ half liquid state. 98

4.1 Pictures show the FTDS- modified surfaces on the quartz crucible. Multiple mini squares of 50x50 μm^2 or big square of 3x3 mm^2. The positioning of a water drop on the hydrophobic surface of the fused silica with the multiple mini squares is also present. 105
4.2 Schematic drawing including the sequences of DX-B tile **a)** Original design. The red strand indicates the strand that will be modified. **b)** DX-B tile after alteration of the middle strand's 5' end (red line) and after modification with the fluorous chain (purple solid line). An example of fluorous chain is pointed out with a purple dashed arrow. 106
4.3 Schematic of the different branch fluorous-phase tags attached to the 5'-end of the DX-B tile DNA strands. **a)** DX-B1 strand with mono-C_8H_{17} and 4 Ts as a linker. **b)** DX-B2 strand with mono-C_8H_{17} 4 Ts as a linker. **c)** DX-B3 strand with mono-C_8H_{17} 4 Ts as a linker. **d)** DX-B3 strand with bis-C_8H_{17} 4 Ts as a linker. **e)** DX-B3 strand with bis-C_8H_{17} 4 Ts as a linker. 108
4.4 1.0% agarose gel of the fluorinated DX-B tiles and DX-arrays. The differences among the samples can be found in Table 4.1. 109
4.5 Schematic shows the immobilisation of the DX-array onto the photolithographically defined fluorous arrays of the fused silica crucibles. DNA strands tagged with fluoro ponytails will strengthen the absorption of the DX-array on the surface. After a simple wash with 50:50 methanol/ buffer, the perfluorinated fused silica crucible is regenerated and ready for re-patterning. 110
4.6 Image of the DX-array with GFP attached (1:1.5 ratio) on the quartz crucible after immobilisation fluorous effect as it was captured using fluorescence microscope (50x objective was used). 111
4.7 **a)** CD and UV absorbance data of DX-array as they were recorded during CD*i* experiment by scanning 4x4 grid array at intervals of 60 μm. **b)** CD and UV absorbance data of the same sample obtained by scanning 14x14 grid array at intervals of 40 μm. 2D maps of SRCD intensity at 220 and 258 nm, respectively, vs x-y; each square represents the spot mapped (40 μm^2) during the 14x14 grid array scan. 113
4.8 2D CD map of single wavelength CD*i* experiment for the DX-array at 265 nm. A 20x20 grid array was scanned of 40 μm step size. 114
4.9 **a)** CD and UV absorbance data of DX-array as they were recorded during CD*i* experiment by scanning 4x4 grid array of 70 μm. **b)** CD and UV absorbance data of the same sample obtained by scanning 10x10 grid array of 60 μm. 2D maps of SRCD intensity at 220 and 258 nm, respectively, vs x-y; each square represents the spot mapped (60 μm^2) during the 10x10 grid array scan. 115

4.10 2D CD map of single wavelength CDi experiment for the DX-array at 220 nm. A 20x20 grid array was scanned of 40 μm step size. 116
4.11 **a)** CD and UV absorbance data of DX-array as they were recorded during CDi experiment by scanning 4x4 grid array of 70 μm. **b)** CD and UV absorbance data generated by DX-array subtraction from the DX-array with Q71 attached sample. 117
4.12 CD data of DX-array with Q71 attached as they were recorded during CDi experiment by scanning 10x10 grid array of 60 μm. 2D maps of SRCD intensity at 208, 220 and 251 nm, respectively, vs x-y; each square represents the spot mapped (60 μm^2) during the 10x10 grid array scan. **b)** Absorbance data of DX-array with Q71 attached 118
4.13 **a)** CD and **b)** absorbance data generated by DX-array subtraction from the DX-array with Q71 attached sample. 2D maps of SRCD intensity at 202, 220 and 251 nm, respectively, vs x-y; each square represents the spot mapped (60 μm^2) during the 10x10 grid array scan. 119
4.14 CD data of DX-array with Q71 attached as they were recorded during CDi experiment by scanning 10x10 grid array of 100 μm using the crucible modified with a big squared pattern of fluorous. 2D maps of SRCD intensity at 208, 220 and 251 nm, respectively, vs x-y; each square represents the spot mapped (100 μm^2) during the 10x10 grid array scan. 120
4.15 CD data generated by DX-array subtraction from the DX-array with Q71 attached sample using the crucible with the mini squares. 2D maps of SRCD intensity at 202, 220 and 251 nm, respectively, vs x-y; each square represents the spot mapped (100 μm^2) during the 10x10 grid array scan . 121
4.16 CD data of DX-array with Q50 attached as they were recorded during CDi experiment by scanning 10x10 grid array of 100 μm using the crucible with the mini perfluorinated squares. 2D map of SRCD intensity at 251 nm, respectively, vs x-y; each square represents the spot mapped (1000 μm^2) during the 10x10 grid array scan. 122
4.17 CD data generated by DX-array subtraction from the DX-array with Q71 attached sample using the crucible with the mini squares. 2D maps of SRCD intensity at 202, 220 and 251 nm, respectively, vs x-y; each square represents the spot mapped (1000 μm^2) during the 10x10 grid array scan. 122
4.18 CD data of DX-array with Q50 as recorded during CDi experiment by scanning 10x10 grid array of 100 μm using the crucible modified with a single square of fluorous modifications. 2D map of SRCD intensity at 220 and 251nm, respectively; each square represents the spot mapped during the 10x10 grid array scan. 123
4.19 CD data generated by DX-array subtraction from the DX-array with Q50 attached sample using the crucible with a big square modified with fluorous. 2D maps of SRCD intensity at 210 and 222nm, respectively; each square represents the spot mapped during the 10x10 grid array scan. . . . 124

1 Mass Spectrometry analysis for CT-A1 ssDNA. 157
2 Mass Spectrometry analysis for CT-A2 ssDNA. 158
3 Mass Spectrometry analysis for CT-A3 ssDNA. 159
4 Mass Spectrometry analysis for CT-A4 ssDNA. 160
5 Mass Spectrometry analysis for CT-A5 ssDNA. 161
6 Mass Spectrometry analysis for CT-A6 ssDNA. 162

7 Mass Spectrometry analysis for CT-A7 ssDNA. 163
8 Mass Spectrometry analysis for CT-A8 ssDNA. 164
9 Mass Spectrometry analysis for CT-A9 ssDNA. 165
10 Mass Spectrometry analysis for CT-A10 ssDNA. 166
11 Mass Spectrometry analysis for CT-B1 ssDNA. 167
12 Mass Spectrometry analysis for CT-B2 ssDNA. 168
13 Mass Spectrometry analysis for CT-B3 ssDNA. 169
14 Mass Spectrometry analysis for CT-B4 ssDNA. 170
15 Mass Spectrometry analysis for CT-B5 ssDNA. 171
16 Mass Spectrometry analysis for CT-B6 ssDNA. 172
17 Mass Spectrometry analysis for CT-B7 ssDNA. 173
18 Mass Spectrometry analysis for CT-B8 ssDNA. 174
19 Mass Spectrometry analysis for CT-B9 ssDNA. 175
20 Mass Spectrometry analysis for the BCN-DNA. 176

21 Mass Spectrometry analysis for DX-A1 ssDNA. 177
22 Mass Spectrometry analysis for DX-A2 ssDNA. 178
23 Mass Spectrometry analysis for DX-A3 ssDNA. 179
24 Mass Spectrometry analysis for DX-A4 ssDNA. 180
25 Mass Spectrometry analysis for DX-A5_5, ssDNA. 181
26 Mass Spectrometry analysis for DX-B1 ssDNA. 182
27 Mass Spectrometry analysis for DX-B2 ssDNA. 183
28 Mass Spectrometry analysis for DX-B3 ssDNA. 184
29 Mass Spectrometry analysis for DX-B4 ssDNA. 185
30 Mass Spectrometry analysis for DX-B5 ssDNA. 186

31 Short axis (SA) cytochrome b_{562} with azF mutation at position 5 and cysteine mutation at position 104. 187
32 Long axis (LA) cytochrome b_{562} with azF position at 50 and cysteine mutatuion at position 21. 188

List of Tables

2.1 The first 3 columns from the left-hand side present the sequences of the oligonucleotides for the CTA tile, whereas the 3 columns from the right-hand side present the sequences of the oligonucleotides for the CTB tile. . 28
2.2 Analytical data of Mass Spectrometry for CTA and CTB. 30
2.3 Sequence of the modified oligonucleotide, where Y is the 5'-Click-easy™ BCN CEP II 32
2.4 Analytical data of Mass Spectrometry of the BCN-DNA. 33

3.1 The first 3 columns from the left-hand side present the sequences of the oligonucleotides for the DX-A tile, whereas the 3 columns from the right-hand side present the sequences of the oligonucleotides for the DX-B tile. 64
3.2 Analytical data of Mass Spectrometry for DX-A and DX-B tiles. 65
3.3 Sequence of the modified oligonucleotide, where Y is the 5'-Click-easy™ BCN CEP II 67

4.1 DX-B tiles and arrays run on the 1.0% agarose gels with the respective fluorous modified strands which are included in each of these samples. . . 109

6.1 Description of **method 1** and solvents used for long DNA strands. 131
6.2 Description of **method 2** and solvents used for short DNA strands. . . . 131
6.3 Required amounts of reagents for the preparation of denaturing urea polyacrylamide gel. 132
6.4 Required amounts of reagents for the preparation of PBS buffer. 134
6.5 Required amounts of reagents for resolving and stacking gel buffers for SDS-PAGE. 134
6.6 Required amounts of reagents for resolving and stacking gel for SDS-PAGE. 134
6.7 Required amounts of reagents for 2 x loading buffer for SDS-PAGE. . . . 135
6.8 Required amounts of reagents for 10 x running buffer for SDS-PAGE. . . 135
6.9 Required amounts of reagents for 10 x TAE/Mg^{2+} buffer for annealing. . 135
6.10 Annealing protocol F1 (90 °C → 4°C). 136
6.11 Annealing protocol F2, second half 50 °C → 4 °C. 136
6.12 Required amounts of reagents for 10 x TAE/Mg^{2+} buffer for annealing. . 137

Abbreviations

A	Adenine
AFM	Atomic Force Microscopy
APS	Ammonium Persulfate
AuNRs	Gold Nanorods
C	Cytosine
CD	Circular Dichroism
CDi	Circular Dichroism imaging
CPG	Controlled Pore Glass
CTA	Cross Tile A
CTB	Cross Tile B
Cytb_{562}	Cytochrome b_{562}
DCM	Dichloromethane
DMT	N,N-Dimethyltryptamine
dsDNA	Double stranded DNA
DX	Double Cross Structure
DNA	Deoxyribonucleic Acid
EDTA	Ethylenediaminetetraacetic Acid
FTIR	Fourier transform infrared
G	Guanine
GFP	Green Fluorescent Protein
GOx	Glucose Oxidase
HCl	Hydrochloric Acid
HRP	Horseradish Peroxidase
HPLC	High Performance Liquid Chromatography
HT	High Tension
LAcytb_{562}	Long Axis Cytochrome b_{562}
LD	Linear Dichroism
MS	Mass Spectrometry
NMR	Nuclear Magnetic Resonance
PAGE	Polyacrylamide Gel Electrophoresis
PBS	Phosphate Buffered Saline (PBS)
PDB	Protein Data Bank
QCM	Quartz Crystals Microbalance
R_F	Perfluorocarbon group
RNA	Ribonucleic Acid
RP	Reverse Phase
SA	Streptavidin
SAcytb_{562}	Short Axis Cytochrome b_{562}
SAG	Substrate-assisted growth
SEM	Scanning Electron Microscopy

sfGFP	Superfolder GFP
SPAAC	Strained ring Promoted Alkyne-Azide Cycloaddition
SPMs	Scanning Probe Microscope
SRCD	Sychrotron Radiation Circular Dichroism
ssDNA	Single Stranded DNA
T	Thymine
TEAA	Triethylamine Acetate
TEM	Transmission Electron Microscopy
TFA	Trifluoroacetic Acid
TFO	Triplex-Forming Oligonucleotides
T_{m}	Melting Temperature
TX	Triple-crossover
SDS	Sodium Dodecyl Sulfate
TEMED	Tetramethylethylenediamine
UV-Vis	Ultraviolet-Visible
VUV	Vacuum Ultraviolter

Nomenclature

$\Delta\varepsilon$	Molar Circular Dichroism
ε	Extinction coefficient
ε_{L}	Molar extinction coefficient for left circularly polarised light
ε_{R}	Molar extinction coefficient for right circularly polarised light
A	Absorbance
c	Concentration
$g - factor$	Anisotropic factor
l	Path length
$mdeg$	Millidegrees
V	Voltage

Declaration of Authorship

I, Christina Xyrafaki , declare that the thesis entitled *Determining Surface Oriented SRCD Signatures of Proteins on DNA Nanostructures* and the work presented in the thesis are both my own, and have been generated by me as the result of my own original research. I confirm that:

- this work was done wholly or mainly while in candidature for a research degree at this University;
- where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
- where I have consulted the published work of others, this is always clearly attributed;
- where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
- I have acknowledged all main sources of help;
- where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
- none of this work has been published before submission

Signed:.....

Date:.....

Acknowledgements

I would like to begin by thanking my supervisor Dr Eugen Stulz for giving me the opportunity to carry out this research, as well as for his endless patience and guidance throughout this challenging journey, and generally for his contribution in my development as a scientist. Many thanks to Dr Rohanah Hussain and Dr Giuliano Siligardi for supporting this project and offering invaluable advice regarding the CD experiments. I would also like to acknowledge Chemistry Department from the University of Southampton and Diamond Light Source for funding my studentship.

Many people have helped me along the way of this PhD. I must thank Dr Dafydd Jones and especially Benjamin Bowen (University of Cardiff) for supplying the modified proteins. I cannot thank Dr David Rusling enough for his advice in the first two vivas and whenever I needed it, for sharing his knowledge on the DX-arrays. Many thanks to Dr Andrea Talandriz Sender, from Glenn Burley's group (University of Strathclyde), for the synthesis of the fluorous tagged oligos, her advice and help with the AFM imaging. I would also like to thank Alasdair Clark and his group (University of Glasgow) for the modification of the quartz crucibles. Many thanks to the B23 team at DLS, who have been very helpful during my beamtime sessions.

I have been so fortunate to work in the same group with Lauren, Ane and Andrea and it was my pleasure to get to know them over the years and share the ups and downs together thoughout this adventure.

But, most importantly I want to thank my family. My dad was defined by his thirst of knowledge, his passion and hard work in everyhting he did and his courage until his last days. I cannot thank him enough for passing these attributes to my sisters and myself. The PhD is not an easy journey, but it is even harder when you are far from your beloved ones- my mum, my sisters, my brothers-in-law and my grandparents have known how difficult this part was for me and their encouragement, trust, love and faith in my abilities were fundamental in order to complete this PhD. Last but not least, I want to thank my partner Theo for being next to me, taking care of me and supporting me.

To the memory of my beloved father

Chapter 1

Introduction

1.1 Proteins

1.1.1 Protein structure

Proteins are the most abundant biological macromolecules occurring in the cells and all parts of cells. Proteins mediate virtually every process that takes place in a cell exhibiting an almost endless diversity of functions, ranging from catalysis to construction. Therefore, proteins are one of the most important molecular instruments through which genetic information is expressed. Their study is vital and covers a vast field in order to understand their structure, functions and impact that conformational changes may have for the protein function and, by extension, the cells.

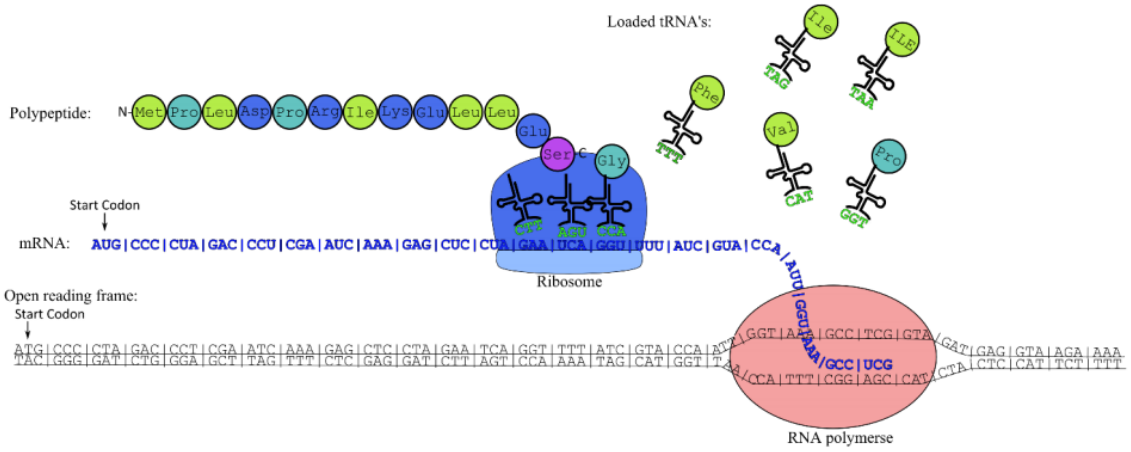

Figure 1.1: Central dogma of molecular biology- protein synthesis. Reprinted from [1]

The protein synthesis that takes place in the cells of all living organisms, consists of two processes, the transcription and translation, which are summed up as central dogma of molecular biology (Figure 1.1).[2] During transcription, DNA is used as a template

to create a molecule of messenger RNA (mRNA) by RNA polymerase. The mRNA then leaves the nucleus, where transcription takes place, and goes to a ribosome in the cytoplasm for the second process, the translation. During this step, the genetic code in mRNA is read and translated to a protein sequence by the ribosome and the transfer RNA (tRNA) in order to make the protein.

As mentioned earlier, there is a great variety of proteins and thousands of different kinds may be found in a single cell. Despite this variety, in humans, all proteins are constructed from the same set of 20 amino acids, the combination of which is coded by the genetic information in the DNA. All 20 amino acids have a carboxyl group and an amino group bonded to the same carbon atom (α-carbon). They differ from each other in their side chains or, as alternatively called *R groups*. The individual amino acids are then joined together with a peptide bond, between the carboxyl group of the first amino acid and the amino group of the next amino acid in the polypeptide chain. This chain, depending on the arrangement of these amino acids, twists into multitude of complex three dimensional shapes and each protein has a specific chemical and structural function, suggesting that each has a unique three-dimensional structure. In general, this is known as *folding* and it is very crucial for the protein's functionality to have the correct folding.[3]

The structure of large molecules such as proteins is commonly defined in four levels, as shown in Figure 1.2. The first, the primary structure, is a description of a unique sequence where all covalent bonds linking the amino acid residues in a polypeptide chain in order to construct each protein. The secondary structure refers to the local three dimensional (3D) shape formed by the folding of the polypeptide chain. The secondary structures are further divided into four main elements such as alpha (α)-helix, beta (β)-strand, turns and unordered. The α-helix structure is a right-handed helical coil with 3.4 amino acid residues per turn, whereas the β-sheet is a ribbon. Intermolecular hydrogen (H) bonding between β-strands forms β-sheets. Furthermore, there is the tertiary level structure of proteins which occurs when the secondary structures of a polypeptide chain are packed together. Finally, when a protein has two or more polypeptide chains (sub-units), their arrangement in space is referred to as *quaternary structure.*[4,5]

Many proteins require covalently bound non-polypeptide units, called *prosthetic groups* for their activity and biological function. The prosthetic groups can either be organic, such as vitamin, sugar groups, lipid, or they can be inorganic such as a metal ion, but none of these groups are composed of amino acids. A well-known prosthetic group is the heme group which is included in hemoglobin and cytochrome c proteins.[7]

Protein function is related to its folding that is determined by biophysical methods. The most popular methods offering high resolution results are the X-ray crystallography, the cryo-electron microscopy and the nuclear magnetic resonance (NMR) spectroscopy.

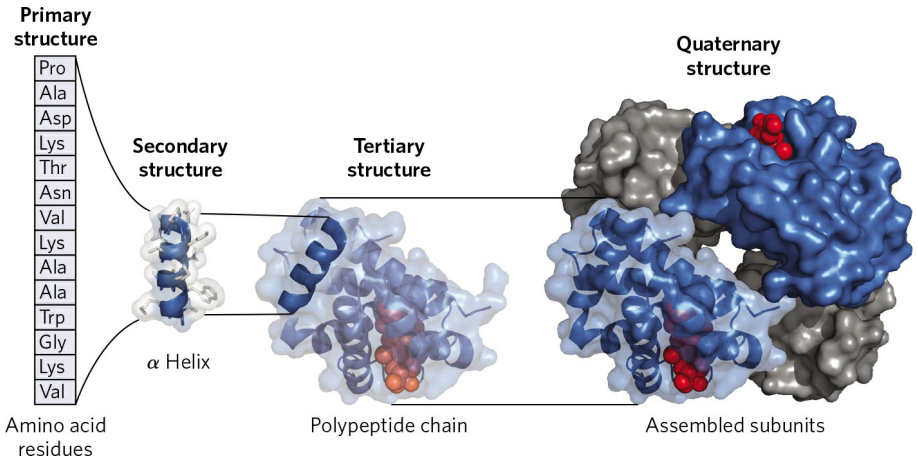

Figure 1.2: Levels of structure in proteins. Reprinted from [6]

X-ray crystallography is a technique used to obtain the three-dimensional structure of a particular protein by x-ray diffraction of its crystallised form. The formation of the crystals creates a lattice in which this technique aligns millions of protein molecules together leading to more sensitive data collection.[8] In addition, the X-ray radiation can often harm or destroy the sample of interest.[9] Although, there are some proteins which can successfully form well-structured crystal lattices, obtaining crystals from other proteins requires more time and effort in order to find the correct conditions of crystallisation. The proteins difficult to crystallise are those with large subunits, high degree of flexibility or high surface hydrophobicity such as membrane proteins.[10] On the other hand, cryo-electron microscopy (cryo-EM) is particularly useful for larger protein complexes, which are difficult to crystallise and the rapid freeze treatment that the sample undertakes maintains its closer-to-native state.[11] Cryo-EM used to have relatively low resolution; however, it has seen rapidly progress in its technical features over the years. Detectors with improved quantum efficiency, as well as the upgrade in the processing algorithms, led to a resolution that is now comparable to crystallography (it reaches 3.5 $\mathring{A}$ for membrane proteins).[12] Furthermore, proteins show multiple conformational states, but only one is reported by crystallography. In contrast, multiple conformations are possible to be observed with cryo-EM, even though this can be challenging for the interpretation of the data.[13] Nevertheless, cryo-EM still requires improvement in order to be able to "compete" the resolution of crystallography, and it is particularly limited to samples of high molecular weight.[14] The third approach is the NMR spectroscopy and it is very important feature since the 3D structure information of macromolecules, such as proteins, can be measured in the natural state (in solution). Furthermore, it is one more technique with high-resolution. However, it is very complicated and difficult to collect information for molecules with large amount of molecular weight and especially since relatively large amounts of pure sample are required in order to achieve reasonable signal.

Other spectroscopic technique used in parallel with the techniques described above, and

which can determine the secondary structure of proteins are the Fourier transform infrared (FTIR) and circular dichroism (CD) spectroscopy which can provide information regarding the overall protein structure. FTIR spectra can be obtained for proteins in a wide range of environments with just a small amount of sample. Nevertheless, the spectra can be complex and sometimes characteristic bands of secondary structure might overlap leading to misreadings. In addition, FTIR is not necessarily protein specific and therefore it can probe every chemical bond that is present in the sample. CD spectroscopy has gained attraction through the years and it can provide qualitative and quantitative information about the protein's secondary structure. There are still some limitations though and similarly to FTIR, only information about the general protein structure can be obtained. However, both FTIR and CD are most powerful when used side by side or in combination with other methods, so that more site-specific details can be collected about the protein of interest. CD spectroscopy is used in this thesis, in combination with Synchrotron Radiation Circular Dichroism (SRCD) spectroscopy, which exploits the high photon flux of the synchrotron beam and the highly collimated micro beamlight leading to higher spatial resolution that the conventional CD. Both methods are discussed in more detail later in this Chapter.

1.1.2 Current studies of proteins on surfaces

The interest in the attachment of proteins and other biomolecules to solid surfaces is growing and is becoming increasingly important in the field of biotechnology. The immobilisation of proteins on surfaces is developed in order to isolate and purify target macro molecules, to perform chemical modifications and enzymatic catalysis, to identify protein-protein, protein-DNA, protein-small molecules interactions and sometimes for diagnostics purposes.[15] Immobilized proteins and enzymes are also used for application in electrochemistry as biosensors for capturing and detecting of macromolecules for analytical purposes. However, further development of these technologies is still required in order to achieve the highest density of protein immobilisation with the minimum amount of protein.

Proteins can be attached to surfaces by (a) physical adsorption, (b) covalent attachment or (c) bioaffinity immobilisation. Physical adsorption does not allow control over the protein's orientation on the surface which leads to randomly oriented proteins and weak attachment. On the other side, covalent attachment is controlled through accessible functional groups of the protein with suitably modified supports, which result in a high surface coverage and irreversible binding. Although bioaffinity immobilisation provides a gentle oriented and homogeneous attachment of proteins like the covalent chemistries, it also offers the ability to detach the protein so that the surface can be reused.[16]

Gold-coated substrates have been widely used as solid surfaces for proteins in different studies. Many studies have focused on covering gold surfaces with monolayers of proteins, such as monolayers of the polypeptide poly (L-lysine).[17] Most of those strategies are challenging, since it is difficult for the proteins to retain their activity and stability. Another approach is to combine proteins with cysteine rich terminal sequences, in order to attach the protein to gold (Au) nanoparticles of various sizes. This motif is believed to provide greatly retained activity of the proteins when they are used as sensor devices by controlling their orientation on the gold surface (Figure 1.3).[18]

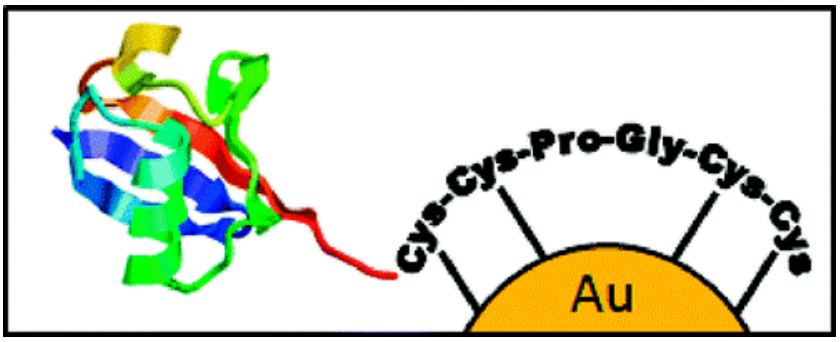

Figure 1.3: Schematic diagram showing the attachment of the protein on the Au nanoparticle. Tetracysteine motif. Reprinted from[18]

Niu *et al.* used the gold surfaces as immunosensor. The gold surface was dipped into a solution of carbon disulphide (CS_2) and protein A, followed by incubation with immunoglobulin G (IgG) solution which formed the sensing surface.[19] The last step was to incubate the sensing surface with anti-IgG, which aims to protect the bioactivity of IgG in the sensor system (Figure 1.4).

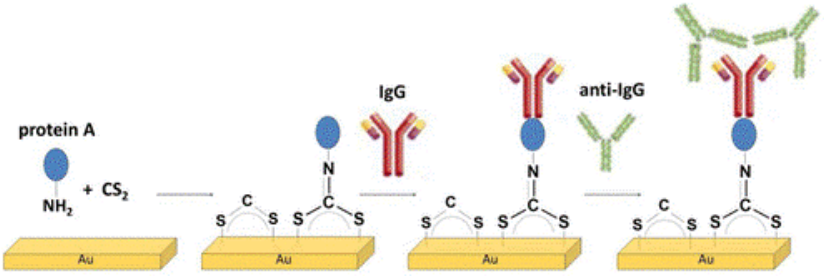

Figure 1.4: Schematic diagram of IgG immobilisation on a gold surface modified with solution of CS_2 and protein A. A gold slide is firstly immersed into the mixed solution of CS_2 and protein A and then, incubated in IgG solution to form the IgG sensing surface. In the final step, the incubation of the sensing surface in anti-IgG solution will allow the evaluation of the bioactivity of the immobilised IgG. Reprinted from[19]

These systems have been complemented with other sensitive techniques for the detection of bimolecular interactions. The binding effects can be monitored by specific signals, and detection systems for the gold modified surfaces include atomic force microscopy (AFM), surface plasmon resonance (SPR)[20,21] and quartz crystals microbalance (QCM).[22]

During the last few decades, DNA nanostructures have also been used as a tool for the attachment of proteins. Park *et al.* tried to create a platform, which would allow the study of macromolecular interactions between interacting proteins when confined in a controlled distance. Their strategy was to attach streptavidin protein arrays with controlled spacing on self-assembled nanostructures.[23] In a similar work from our group, proteins such as sfGFP and β-lactamase, were modified with one single-stranded DNA (ssDNA) at a precise location to achieve a well-defined orientation on DNA origami tiles, in order to create a system that allows the study of the impact of orientation on protein function.[24] This research project is based on this approach, but a more extended DNA nanostructure is chosen.

1.2 Deoxyribonucleic Acid, DNA

1.2.1 Building blocks of nucleic acids and tertiary structure of DNA

The history of the deoxyribonucleic acid (DNA) is commonly associated with the names of Watson and Crick, who deciphered the structure of the DNA molecule as a three-dimensional double helix in 1953.[25,26] DNA consists of two antiparallel single strands, which contain several subunits called *nucleotides* (Figure 1.5 (a)). The nucleotide is composed of a phosphate group, a 2'-deoxyribose (pentose sugars) and one of four heterocyclic nucleobases, linked together. The four nucleobases which appear in DNA are adenine (A), guanine (G), cytosine (C) and thymine (T), which are shown in Figure 1.5 (b).[26] Each of these nucleobases is based upon the structure of either purines (A and G) or pyrimidines (C and T). The nucleobases are attached to a sugar ring forming a nucleoside. Adding a 5'-phosphate group to the nucleoside forms the nucleotide and finally, when different nucleotides are linked together by a 3',5'-phosphodiester bond, they form a linear polymer.[27]

The two antiparallel strands are held together by the bonds between opposite bases, pointing inside the helix with phosphoester backbone on the outside.[28] These bases bond to one another by hydrogen bonds, following the Watson-Crick base pairing: A-T and C-G. A-T has two hydrogen bonds, whereas C-G has three hydrogen bonds and all bonds are depicted in Figure 1.6.

Over the years, it was found that DNA actually is a polymorphic system, which can adopt more geometries than just the well-known double helix. This structural polymorphism depends on various factors such as the surrounding environment, the nucleotide sequence and the type of base pairing.[29,30] Therefore, DNA exists in many possible conformations, but three of them are the main forms: A-form, B-form, Z-form, as shown in Figure 1.7. Both the A-form and B-form are right-handed helices while the Z-form is a left-handed helix. The B-DNA is the most common form and this is the type naturally occurring

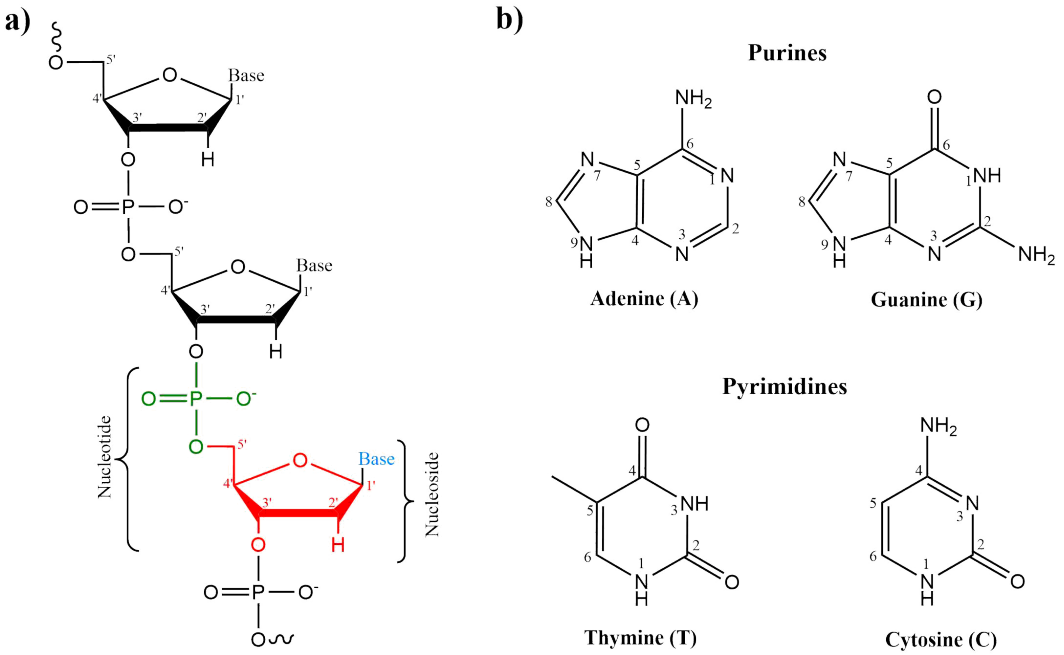

Figure 1.5: a) The building blocks of nucleic acids: phosphate group (green), sugar unit (red), nucleobases(blue). **b)** Structure of four nucleobases.

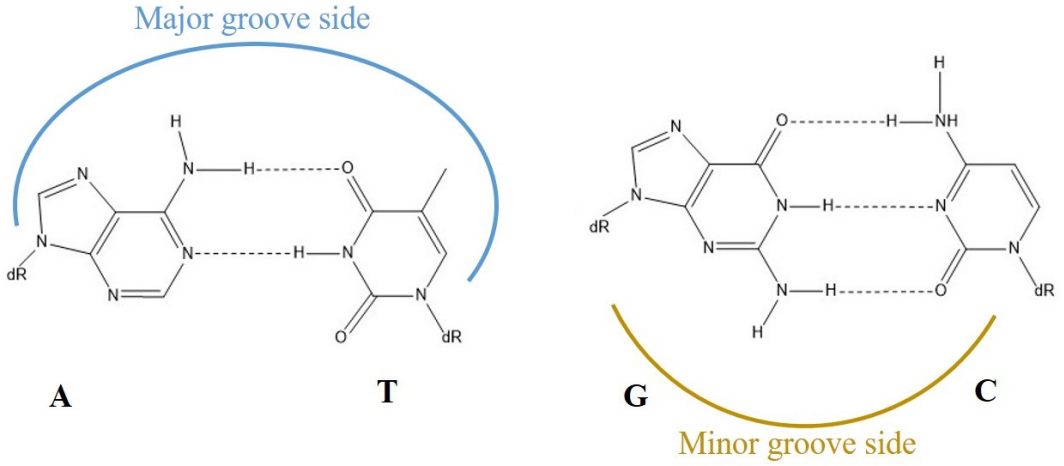

Figure 1.6: Watson-Crick Base pairings A-T and C-G. dR indicates the position of the deoxyribose.

in humans. The A-DNA is mainly found in RNA-RNA and RNA-DNA duplexes and it can also occur under non-physiological conditions in partly dehydrated samples of DNA, whereas the Z-DNA is unfavourable but it can be formed by stretches of alternating purines and pyrimidines, e.g. GCGC, in a solution of high salt concentration.

The B-DNA is not only recognised as a biological material, but also as a nanomaterial and because of its geometry and its helical parameters (Figure 1.8) it has interesting applications in technologies based on nucleic acids. Its appealing properties for use in nanotechnology are: (1) its minuscule size with a diameter of 2 nm, (2) the short structural repeat, known as *helical pitch*, of 3.4-3.6 nm, and (3) its stiffness with a

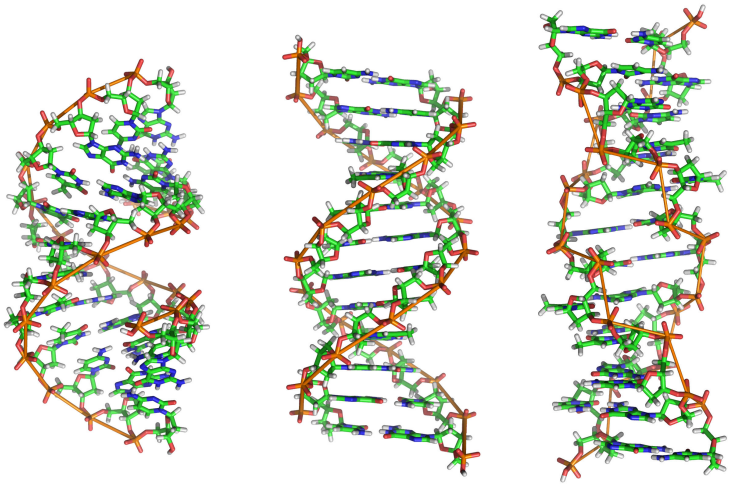

Figure 1.7: Three main forms of DNA, from left to right: A-DNA, B-DNA, Z-DNA. Reprinted from[31]

persistence length of around 50 nm (150 base pairs).[32] The DNA as a nanomaterial will be further discussed in Section 1.3.

Parameter	A-DNA	B-DNA	Z-DNA
Helix sense	Right	Right	Left
Base pairs per turn	11	10	12
Axial rise (nm)	0.26	0.34	0.45
Helix pitch (°)	28	34	45
Base pair tilt (°)	20	–6	7
Twist angle (°)	33	36	–30
Diameter of helix (nm)	2.3	2.0	1.8

Figure 1.8: Different helical parameters for the common DNA conformations. Reprinted from[33]

The tertiary arrangement of nucleic acids is not limited to a double helical geometry. In particular, constructions like triplex-DNA[34] (three-stranded DNA or also known as H-DNA) and G-quadruplexes[35] are found. Triplex-DNA is thought to strongly influence the DNA transcription processes, and the expression of normal and disease-linked genes.[36,37] G-quadruplexes seem to be involved in multiple cellular processes, such as transcription, translation, replication and the regulation of telomere structure.[38] This structural overview is not conclusive, but it should be clear that the shape diversity of nucleic acids is important for them to maintain their biological function.

1.2.2 Oligonucleotide Synthesis

Solid phase synthesis is the most common way to synthesise DNA. It contains a number of reactions that builds up the DNA from the 3'-terminus to the 5'-terminus (Figure 1.9). A solid support bead is used for the reaction, which is made from either controlled

pore glass or polystyrene. Some beads contain the first base of the sequence bound via an ester linkage and some "blank" beads allow the modification of the 3'-terminus.

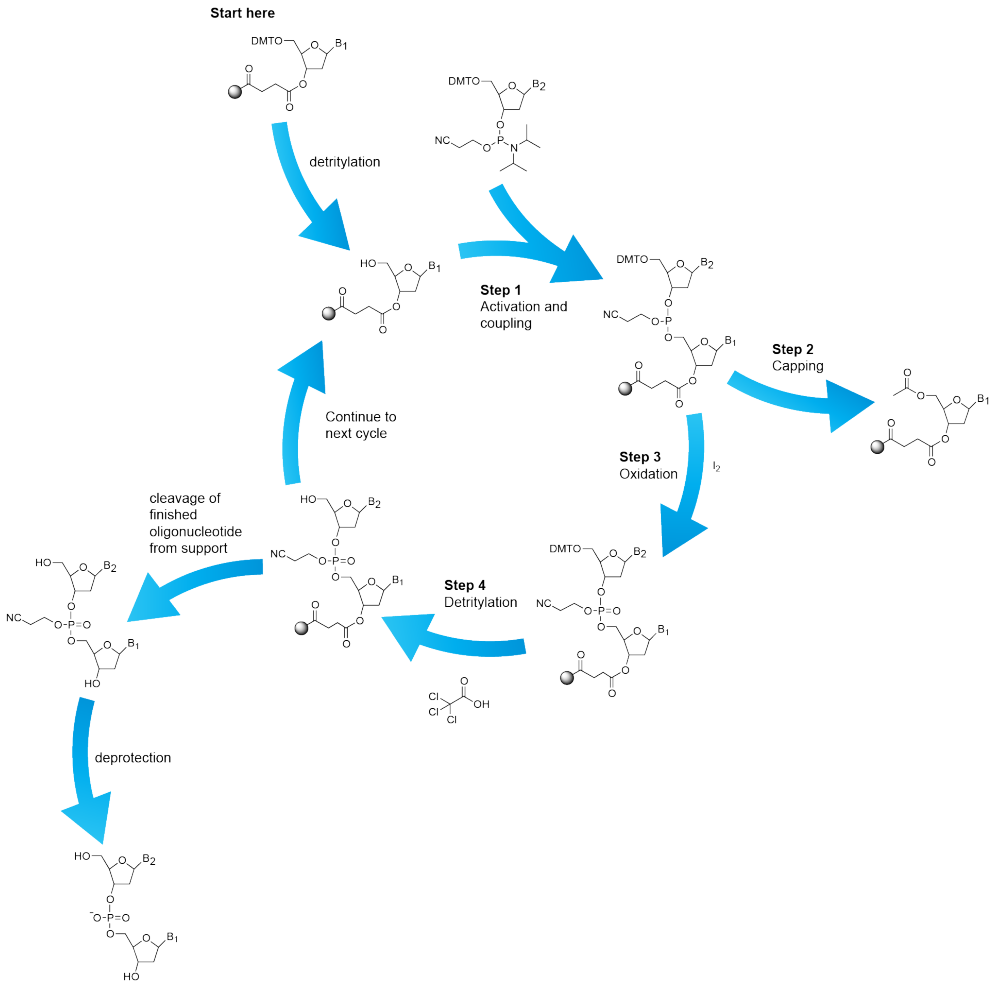

Figure 1.9: The phosphoramidite oligonucleotide synthesis cycle (ATDBio Nucleic Acid book). Reprinted from[39]

During the first step, the dimethyoxytrityl (DMT) protection group of the 5'-oxygen is removed using dilute trichloroacetic acid. An orange colour is visible from the cleaved trityl group, which can be used to estimate the coupling efficiency by monitoring its absorbance. The deprotected hydroxyl group is now free to react with the next phosphoramidite monomer.

The nucleoside phosphoramidite is activated by an activator (tetrazole) and both of them are dissolved in acetonitrile. The diispropylamino group of the nucleoside phosphoramidite is protonated from the tetrazole catalyst and this makes it a good leaving group. Therefore, the leaving group is displaced by the 5'-hydroxyl group, forming a new oxygen-phosphorus bond and creating a support-bound phosphite triester.

A small percentage of unreacted 5'-hydroxyl groups are blocked during the capping step by acetylation. The unreacted alcohol groups are then unable to react to further monomer additions.

Subsequently, the phosphite triester [P (III)] is unstable in acidic conditions and needs to be oxidised to an oxidation state P (V), which is more stable for the next steps of the cycle. This step is achieved by iodine oxidation. Another detritylation of the DMT hydroxyl group on the 5'-monomer begins and the cycle is repeated, once for each base, to produce the required oligonucleotide. Finally, when the final oligonucleotide is ready, it is cleaved from the solid support (beads) using ammonium hydroxyde, which is also used for the deprotection of the groups that are protected on the bases.

1.3 DNA as structural material

Today, DNA is considered far more than just a molecule and it has become the icon of modern biosciences. Therefore, remarkable progress is made to further understand the human genome and the impact of DNA to life and health. By taking advantage of its base-pairing rules and structural characteristics, DNA is able to spontaneously form into larger structures, which has led to the development of research area known as *bio-nanotechnology.*[40] Researchers working in DNA bio-nanotechnology made a great stride in creating building blocks or DNA tiles with nanometre precision.[41]

The field of DNA bio-nanotechnology was pioneered by Ned Seeman, who created various important DNA motifs for self-assembly, such as Holiday Junctions (Figure 1.10).[32,42] The four Holiday branched junctions include protruding complementary sticky ends which allow for the self-assembly of large arrangements of DNA.

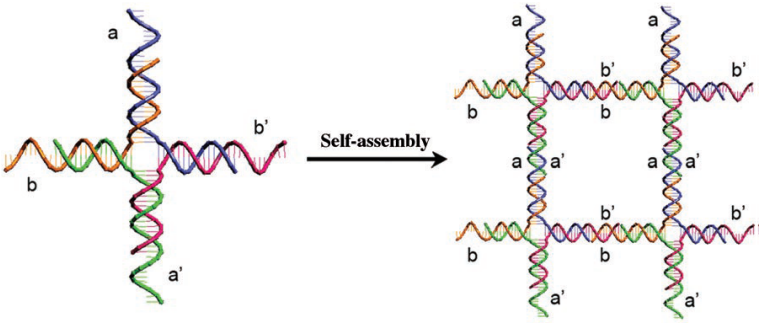

Figure 1.10: Four arm-DNA branch junctions with complementary sticky ends (a-a' and b-b' pairing) to self-assemble into 2D lattice. Reprinted from[42]

This is the fundamental principle behind DNA nanostructure assembly *via* the tile-based method and the beginning for numerous kinds of DNA nanostructures to be developed, including double-crossover (DX) DNA tiles, triple-crossover (TX) tiles, 4 x 4 tiles, and three-point-star tiles.[32,43–45] These nanostructures are shown in Figure 1.11. The tiles can be combined in different ways using sticky ends, in order to create higher order nanostructures such as 2D lattices, nanotubes and more complicated 3D structures.[46]

Paul Rothemund proposed the technique of folding DNA to build complex shapes and patterns, which are referred to *"multi-stranded"*, *"scaffolded"* and *"single-stranded"*,

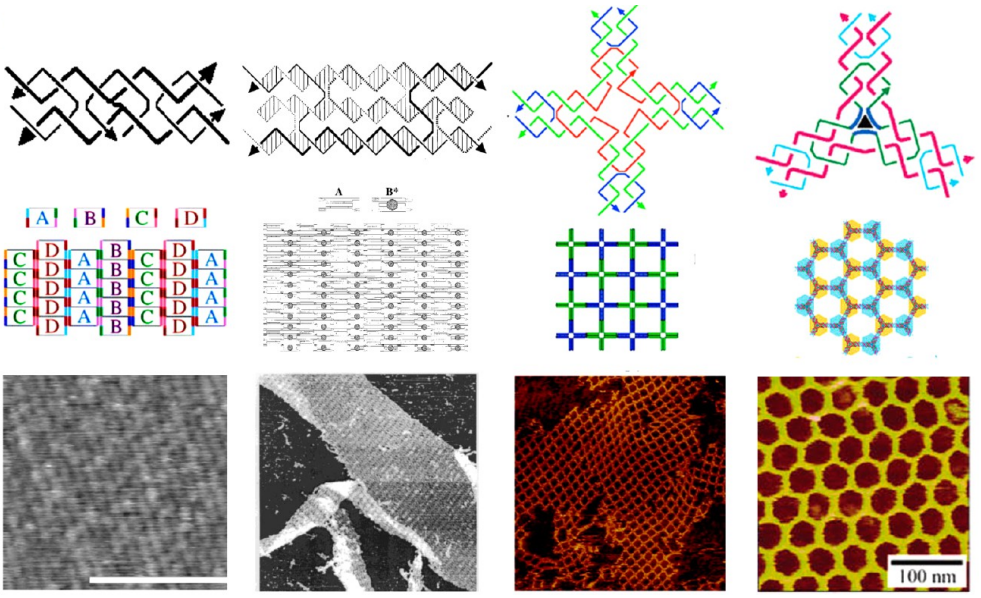

Figure 1.11: From left to right: double-crossover (DX) tiles, triple-crossover (TX) tiles, 4x4 tile and three-point-star tile. The top images are the structures of motifs, the middle images 2D and 3D assembly patterns, and the bottom images are AFM images of the lattices. Adapted from [46]

with the last two known as *DNA origami*.[47] The scaffolded DNA origami structure is usually composed of one long strand of circular virus DNA, called *scaffold strand*, and a set of shorter oligonucleotide strands, called *staple strands*, which finally fold into designed 2D shapes. Each of the typically ∼200 staple strands are designed to be complementary to multiple regions on the long scaffold strand. The most commonly used scaffold is M13mp18, a single stranded, 7249 nucleotide, and circular DNA genome derived from bacteria phage M13.[48] Exploiting the same long virus genome strands and different staple strands, many different 2D shapes can be created such as squares, rectangles, five-pointed stars, smiley faces and different triangles with approximate dimensions of 70 nm x 100 nm (Figure 1.12).[49] 3D structures have also been achieved by Douglas *et al.*, which could be bent and twisted following established design rules to give structures like rugby balls and nanoflasks.[50,51]

1.3.1 DNA nanostructures as templates for self-assembly

DNA origami, and in general DNA nanostructures, became highly versatile tools with many applications relying on the incorporation of guest molecules into these structures, exploiting the DNA as a scaffold or template in order to control their spatial positioning. The two-dimensional DNA nanostructures can contain specifically addressable sequences to which further functionalities can be attached. Diverse range of molecules including nanoparticle, proteins or enzymes, antibodies and other chemical groups have been incorporated in these nanostructures through the years.

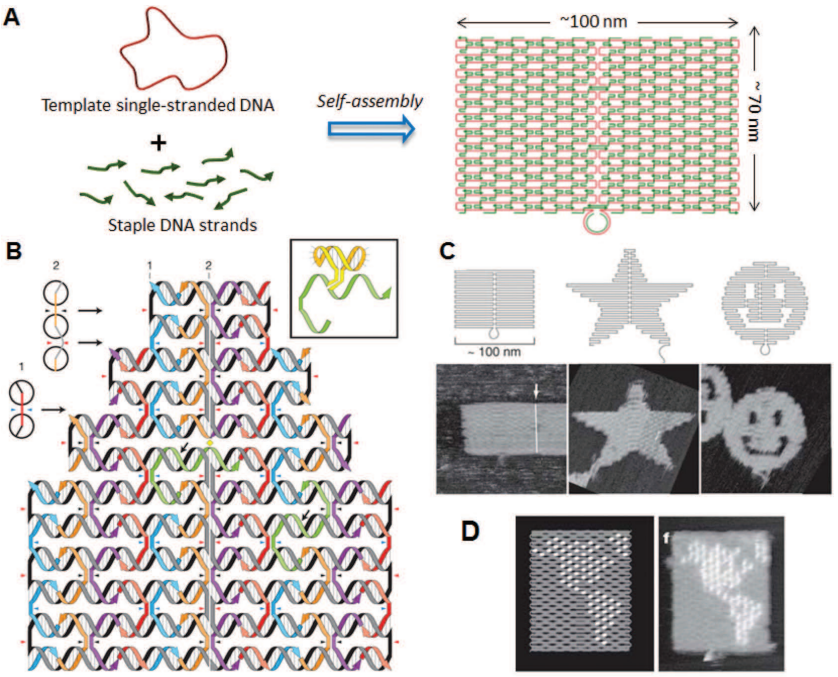

Figure 1.12: DNA origami structure. **(A)** Method employed to prepare a DNA origami structure. **(B)** Design of a self-assembled DNA origami structure and geometry of the incorporated dsDNAs. **(C)** Design and AFM images of self-assembled DNA origami structures. Adapted from [49]

Zhang *et al.* used the 2D DNA nanogrids to arrange Au nanoparticles with accurate control of interparticle spacing in order to be used in nanoelectric and nanophotonic devices.[52] In 2002, Niemeyer *et al.* proposed the organised attachment of Flavin mononucleotide (NAD(P)H:FMN) oxidoreductase and luciferase with a single stranded DNA template.[53] Based on this work, Engelen *et al.* used two enzyme systems, the horseradish peroxidase (HRP) and the glucose oxidase (GOx), which were assembled with DNA nanostructures. The goal of this study was to investigate the enzyme activity and it was concluded that the best possible rates are when the interenzyme distance is reduced, and the proteins are actually in touching distance.[54] The programmability of the DNA nanostructure was also exploited by Selmi *et al.* who used the 2D DNA templates to attach proteins with the aim of creating non-overlapping arrays of protein molecules suitable for imaging by cryo-electron microscopy. This technique could be beneficial for studies of hard-to-crystallise membrane proteins or other protein complexes.[55]

As mentioned earlier in Section 1.1.2, recent work from Stulz's group used DNA origami tiles for attachment of proteins on a well-defined orientation in order to create a system that can allow the study of the impact of orientation on protein or enzyme function.[24] In this thesis, this application is taken a step forward in order to attach proteins in more extended DNA nanostructures, such as the 2D DNA nanogrid or the DX-array, and study the protein orientation and, by extension, their structure using SRCD spectroscopy.

1.4 Protein-DNA conjugation

DNA nanostructures are a tool with multiple uses, since it is a well-defined template. One of their most interesting attributes is the ability to assemble proteins on DNA tiles at precise locations. To achieve the precise assembly of the protein to the DNA origami structures, a method of conjugation of protein and DNA nanostructure is required. Proteins are made up of amino acids and they are folded into a three dimensional structure. The functional groups of the amino acids, which are on the surface of the proteins, are ideal targets for the conjugation with the DNA templates without any enzyme modification. Cysteine and lysine sites are usually used as targets, where the ssDNA is conjugated to the protein through cross-linking agents to form covalent linkages between the protein and the DNA. This method can also be used for the attachment of proteins on other solid surfaces, such as gold, as mentioned earlier. The disadvantage of this approach is that there is usually more than one attachment points on the surface of the protein leading to ill-defined conjugates with potentially more than one DNA strand attached to the protein.[56–60] Other methodologies that exist are the biotin-streptavidin coupling, formation of fusion proteins or the addition of terminal sequences tags, but unfortunately all of them restrict the attachment site of the modifier.[61–64] Some examples of these methodologies are presented in Figure 1.13.

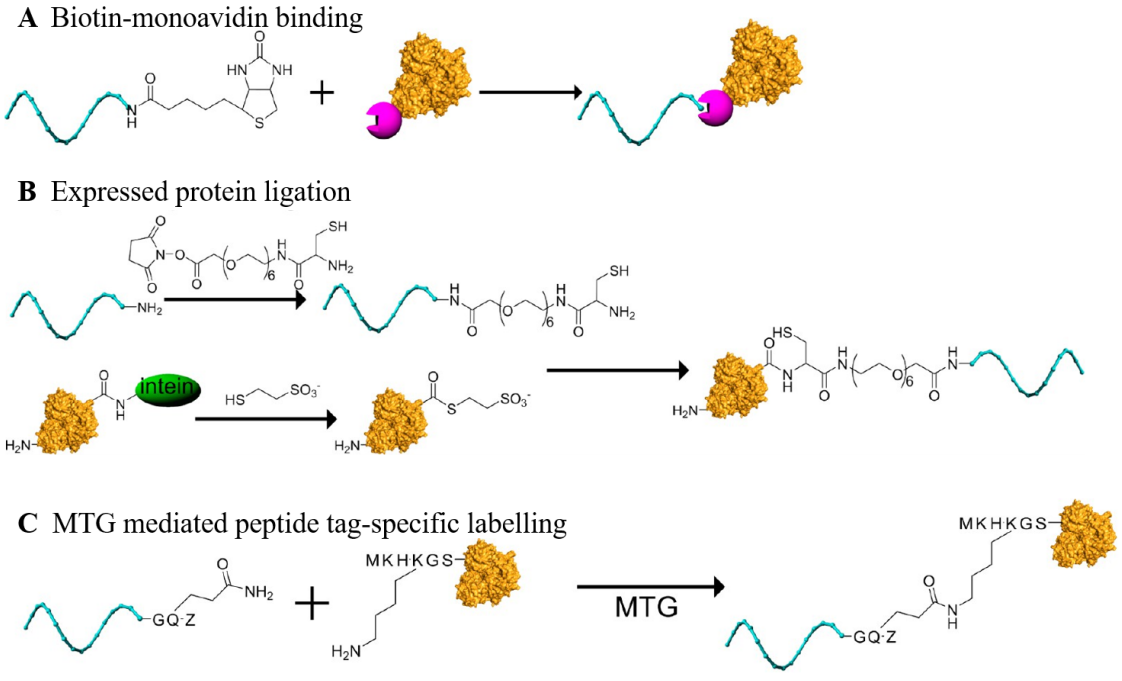

Figure 1.13: **(A)** Biotin-monoavidin binding. **(B)** C-terminal fused intein protein-DNA conjugation by expressed protein ligation. **(C)** Acyl transfer reaction catalysed by microbial tansglutaminase (MTG) between N-carbo benzyloxyglutaminylglycine (Z-QG) DNA and short peptide tag. The schemes were adapted from[46]

Therefore, it is important to have a well-defined protein-DNA conjugation in order to control the orientation of proteins on the DNA nanostructure's surface. One suggestion

to this problem was the use of specific cysteine residues. Nevertheless, the structure of the protein might be affected.[65] A more interesting perspective is to involve specific modification of the proteins regardless of the original amino acid structure. Duckworth *et al.* performed a copper catalysed alkyne-azide Click reaction and used C-terminal tetrapeptide tagged proteins with azide-modified isoprenoid diphosphate.[66] Yet, this may induce further problems since CuAAC can affect the function of the protein by changing its structure.[67–69]

In Marth's work, the approach that was followed and is followed in this research project as well, is the introduction of a non-biological reaction handle at a specific point in a protein, such as an azido-modified amino acid through engineering of the genetic code. In particular, azido-phenylalanine is introduced into proteins at specific points. This reaction handle can be used for the attachment of the ssDNA through a copper-free strained ring promoted alkyne-azide cycloaddition (SPAAC) reaction (Figure 1.14.[24] The attachment points of proteins are precisely defined through genomic engineering and hence, the control and accuracy of the assembly can be tuned. By choosing an appropriate residue, ssDNA conjugation had minimal impact on protein function, even when attached close to active sites. This design approach allows fine-tuning of protein assemblies on DNA nanostructures in order to precisely study protein activity and therefore, this will be vital for the biological systems where orientation is essential, such as membrane bound proteins or in multienzyme assemblies. It should, however, be noted that the approach of using genomic engineering can induce the same issues as noted above: the introduction of non-natural amino acids may well have an impact on the protein structure, which will need to be evaluated, for example using CD spectroscopy by comparing the native vs mutated protein.

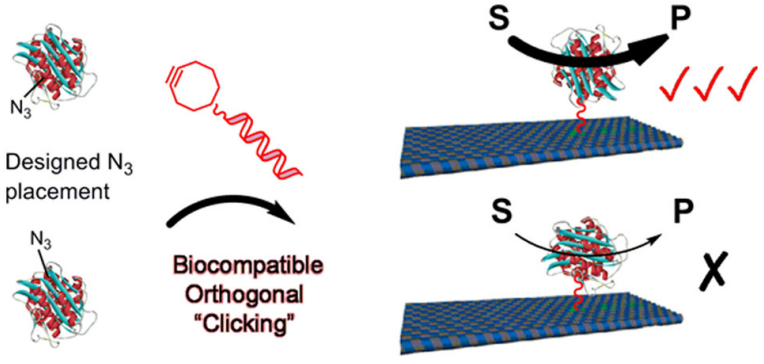

Figure 1.14: Schematic of the genetically encoded β-lactamase with the azide in different residues and its further modification with the ssDNA through biocompatible and orthogonal click chemistry. Finally its attachment to the DNA origami tiles. The conjugation position and by extension the orientation of β-lactamase to the DNA origami tile is important in terms of catalytic activity. In the case at the top, this conjugation position leads to greatly improved enzymatic activity than the free enzyme and in the case at the bottom, it shows lower activity and therefore, product transformation. Reprinted from[24]

Marth's approach was part of preliminary work to this thesis by exploring click chemistry to enzymes, which were then attached to DNA origami tiles, in order to measure the activity and see the impact that the genomic engineering and the different orientations of the enzyme on the tile may have to its general activity. In this thesis, this approach has been taken a step forward by applying it in different systems, more extended and well-oriented surfaces of DNA nanostructures, in order to obtain oriented CD spectra of proteins or enzymes and correlate the spectral features to specific orientation.

1.5 Introduction to circular dichroism and synchrotron radiation circular dichroism

1.5.1 Circular Dichroism (CD) and applications

Nature is remarkably skilled at creating and tailoring chirality of biomolecules in order to accomplish specific functions, such as chiral recognition in biochemical reactions. The majority of biomolecules (like proteins, nucleic acids, lipids, and carbohydrates) and modern drugs are chiral and their chirality is conveniently tested by measuring Circular Dichroism responses. The phenomenon Circular Dichroism (CD) refers to the differential absorption of left- and right- handed circularly polarised light passing through chiral molecules.[70,71] For a given wavelength the molar CD ($\Delta\varepsilon$) is given by:

$$\Delta\varepsilon = \varepsilon_{\mathrm{L}} - \varepsilon_{\mathrm{R}} \tag{1.1}$$

where ε_{L} and ε_{R} are the molar extinction coefficients for left and right handed circularly polarised light, respectively. After the circularly polarised light passes through an optically active sample, both components are absorbed to a different extent, resulting in an elliptical polarisation, which is measured by a spectropolarimeter and referred to as the *CD effect* and initially called the *Cotton effect.*[72] CD is a widespread and well-established spectroscopic technique in structural biology alongside with other high-resolution techniques such as X-ray crystallography and NMR spectroscopy. Despite this, for structural studies of proteins, such as membrane proteins, the NMR spectroscopy and X-ray crystallography are rather challenging, due to their size or irregular structure and their difficulty to crystallise.[73] Although CD is a low-resolution method, it is also a fast and easy spectroscopic technique to study the structural behaviour of proteins and other chiral bio-macromolecules in their native environment, i.e. aqueous or membrane-bound, though the latter is proving difficult due to the experimental setup and stability of the proteins. The main issues arise from overexpression of the proteins, stabilising in suitable media, and high background signals of lipids. CD is finally considered as a sensitive physical technique to determine structures and monitor structural

changes, since in the case of proteins, it can directly interpret the changes of protein secondary structure, even though the method is empirical.[74]

Information about the secondary structure of proteins can be obtained from the CD spectra in the far UV wavelength region, which is between 180 and 250 nm, and the near-UV spectra reflect the contributions of aromatic side chains, disulfide bonds, and induced CD bands of prosthetic groups. The most studied CD signatures of proteins are the secondary structural elements, such as the alpha-helix, the beta-strand, turns and unordered. The far UV spectra of these groups are shown in Figure 1.15. Apart from the information about the secondary structure of proteins, the ligand-protein interactions can also be determined, and the folding and unfolding of the protein can be observed through changes in the CD spectrum.[75] Each band in the spectra is associated with an electron transition of the peptide bond. The spectra of the proteins in the far UV regions are dominated by the n→π* and π→π* transitions of amide groups, and they are influenced by the geometries of the polypeptide backbones.[76] At these wavelengths, the chromophore is the peptide bond and the signal is present when it is located in a regular, folded environment. The first and the weakest energy transition in the peptide chromophore is an n→π* transition observed at 210-220 nm, which involves non-bonding electrons if the oxygen of the carbonyl. On the other hand, the strongest one is an absorption band at 190 nm deriving from π→π* transition involved the π-electrons of the carbonyl.[77,78]

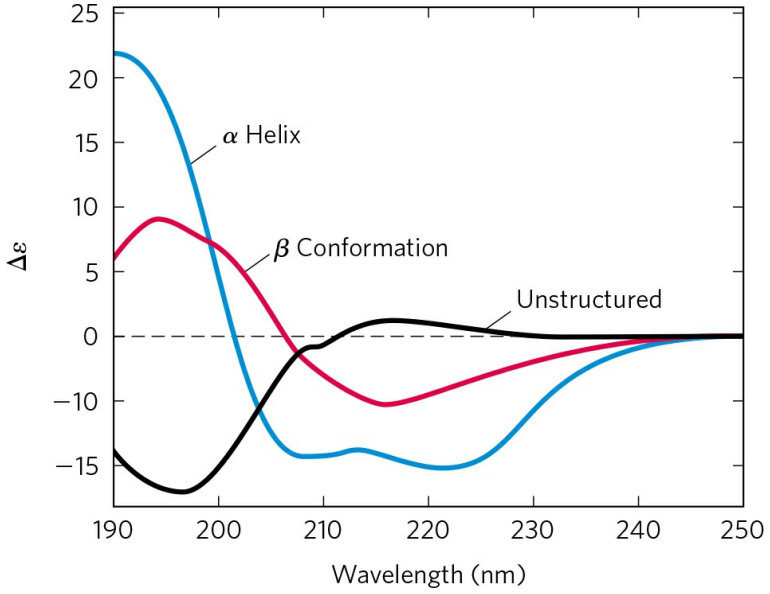

Figure 1.15: CD signatures of α-helical protein (blue), β-strand (red) and unordered (black) proteins. Reprinted from[7]

As determined and according to Figure 1.15, in the case of the unordered proteins, the far UV CD is positive at 212 nm (n→π*) and negative at 195 nm (π→π*). For the β-strand, there is a negative peak at 218 nm (π→π) and a positive at 196 nm (n→π*). However,for the α-helix, the 208 nm negative band and the 190 nm positive band result

from exciton splitting of the $\pi \rightarrow \pi^*$ absorption band into a long-wavelength component polarised perpendicular to the helix axis (190 nm) as predicted by Moffitt.[79] Finally, there is also a negative band with a maximum at 222 nm which is assigned to the $n \rightarrow \pi^*$transitions.[80]

CD spectroscopy is widely used for determining the secondary structural content and integrity of proteins but, it is very useful to elucidate the electron structure of DNA, which is linked to the conformational structure. The four bases of DNA account for optical transitions in the UV due to their delocalised π-electrons. However, they are planar and therefore, they are not chiral and alone they do not give CD signal.[81] On the other hand, the sugars of nucleosides have an intrinsic asymmetry, and therefore, they are chiral and they induce a small CD into the electronic transitions of the bases in the monomeric nucleotides and induce the superasymmetric helicity of the oligomers and polymers.[74] The interaction of the strong $\pi \rightarrow \pi^*$ transitions of the chromophoric bases with the high energy in the sugars lead to a CD of low intensity.[82] However, there are also two types of $n \rightarrow \pi^*$ transitions which can take place at the same wavelength region. The first one corresponds to non-bonding electrons on the nitrogens of the bases that are excited to empty π^* orbitals, and those transitions have moderate intensity. The second $n \rightarrow \pi^*$ transition is due to the non-bonding electrons on the amidelike oxygens which can also be excited to π^* orbitals but those are of low intensity. Each CD band corresponds to a transition in the normal absorption, so some of the extra CD bands may well be due to underlying $n \rightarrow \pi^*$ transitions which are though buried under the intense $\pi \rightarrow \pi^*$ transitions.[74] In fact, circular dichroism of nucleic acids is mainly dependent on the stacking geometry of the bases. The hydrophobic stacking of the bases in the oligomers and polymers leads to close contact and coulombic interactions that result to intense CD bands which correspond to each base transition. The details of the CD spectrum are determined by the nature of the interactions and that makes CD spectroscopy extremely sensitive to secondary structure.[83]

CD is a valuable technique because nucleic acids are polymorphic, and they can assume a variety of secondary structures. As it was discussed in Section 1.2, DNA can assume many conformations that depend on the salt, solvent, humidity and base composition. Each conformation has a characteristic CD spectrum that depends on somewhat on the sugar type and base composition. In Figure 1.16, the CD spectra of A-RNA, A-DNA, B-DNA and Z-DNA are presented. It is evident that each conformation has a very characteristic CD signal. The A-RNA and A-DNA have similar CD in terms of shape, with a difference in the positive peak, which is shifted for the A-DNA at 270 nm instead of 260 nm that is seen for the A-RNA, and furthermore, in the intensity of the negative peak at 210 nm.[84,85] The B-DNA shows approximately equal maximum and minimum at 275 nm and 245 nm, respectively, and a crossover around 260 nm.[86] The positive peak of B-DNA gives less than one-half the magnitude of A-DNA positive peak. In fact, the shape and magnitude of CD spectra depend on the base sequence,

but that does not mean that the overall patterns do not remain constant.[87] On the other hand, the Z-DNA has a different behaviour with conservative spectra at 240 nm, a negative peak at 290 nm and a positive one at 260 nm, showing that the CD will also depend on the orientation.[88] According to literature and the CD data that have been collected using vacuum CD, CD is a useful technique to compare conformation of DNA or RNA and differentiate right-handed double helix or nucleic acids from left-handed. The former shows a positive peak around 186 nm whereas the latter gives a negative peak.[88] Therefore, a good method to establish the sense of a duplex helix is to measure the CD in the wavelength range from 170 to 220 nm.

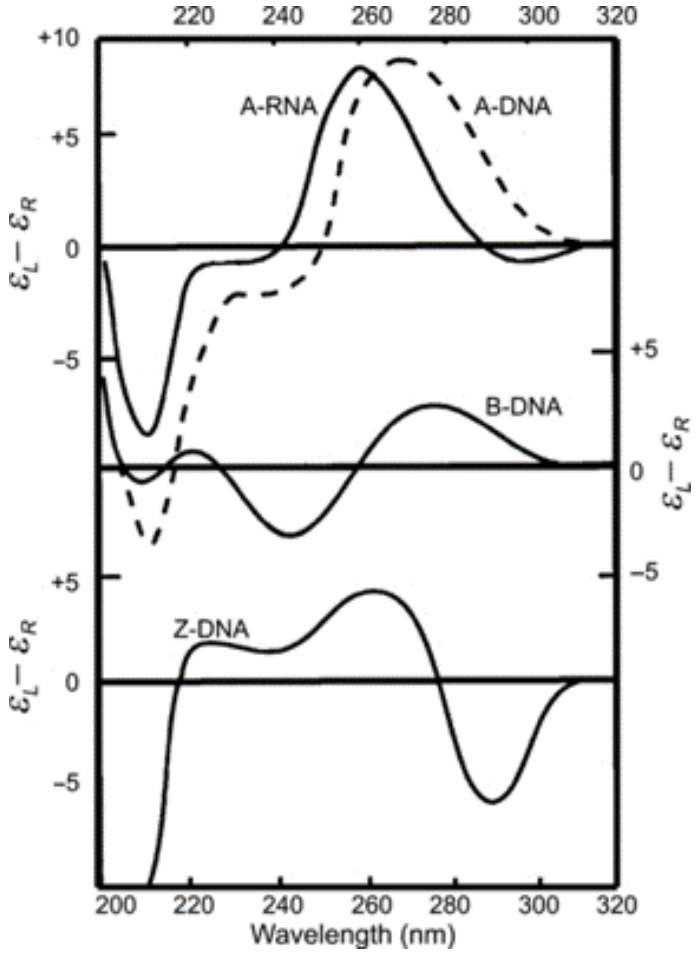

Figure 1.16: Circular dichroism spectra for right-handed A-RNA and A-DNA, right-handed B-DNA, and left-handed Z-DNA. Adapted from[85]

Many studies have also been made on CD characteristics of natural molecules, such as peptides and double-stranded DNA, when used as templates for the growth of chiral metal (plasmonic) nanostructures.[89–92] Those plasmonic nanostructures have several applications in optics, sensing, medical diagnostics and therapeutics.[92,93] CD spectroscopy is a useful tool in order to explore the new CD response at the plasmon resonance, generated by the achiral plasmonic nanoparticles and the chiral molecules (peptides, DNA), the CD intensity of the chiral molecules in the UV region, and the plasmon-plasmon

interactions between the building blocks.[93] Figure 1.17 shows, characteristic example of CD studies of plasmonic nanostructures.

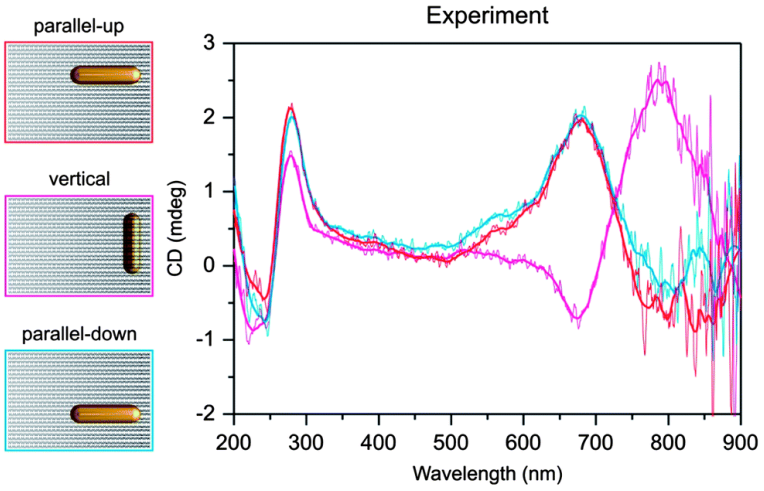

Figure 1.17: Optical characterisation of the DNA origami dressed with single AuNRs. Chiral molecule-induced CD is clearly observable centred at around 730 nm. CD in the parallel AuNR cases (red and blue) shows opposite chirality compared to that in the vertical AuNR case (purple). The experimental results clearly demonstrate that the chiral molecule induced CD is highly orientation dependent. Reprinted from [92]

CD spectroscopy can also be used for other applications, such as monitoring structural changes. For instance, a study on lipid membrane anchors, which are suitable for solid-supported DNA synthesis, used the CD method in order to examine whether the duplex structure of the modified dsDNA is distorted after thermal denaturation experiments.[94]

The bench-top CD instruments have been improved in the past decade and CD spectra can be measured in the far UV region down to 180-185 nm. Nevertheless, the vacuum ultraviolet (VUV) region cannot be reached with these instruments, since there is a limit of signal detection for those using Xe-lamp light source. It is very important to be able to measure CD spectra at least down to 185 nm as important CD spectral feature at about 190 nm can be very diagnostic of the protein-DNA template folding.[95]

1.5.2 Synchrotron Radiation Circular Dichroism (SRCD) Spectroscopy

Synchrotron Radiation Circular Dichroism (SRCD) is a rapidly growing technique used as a light source for CD, which is referred to as *SRCD spectroscopy*. SRCD has extremely evolved over the years for the structural analysis of proteins and other chiral biomaterials and bio-macromolecules.[95,96] In comparison with the bench-top CD, the high photon flux of a synchrotron beam does not have the limitations of the Xe-lamp, and therefore it extends the wavelength range to the VUV region, often down to 130 nm, while providing an extremely bright light source. Thus, only small amounts of biological samples are required and the spectra allow for a more accurate secondary structures estimation of proteins. In addition, even with the use of strongly absorbing matrices, such as buffers,

salts or scattering lipid vesicles, the measurements of the protein have high quality because very short path lengths can be used which suppress scattering and absorbance of salts and buffers.[97]

The SRCD beamline B23 at Diamond Light Source, UK, provides a unique highly collimated microbeam sine qua non for CD *imaging* at the highest spatial resolution. Unlike other existing synchrotron beamline for SRCD, B23 beamline can measure a broad wavelength range from the VUV to the visible spectral region (140-650 nm), which is important since the absorption bands of the peptides or the aromatic amino acids of the proteins absorb in this range.[96] B23 beamline is used for SRCD spectroscopy and it can overcome most problems in the study of proteins, and especially of the membrane proteins that are available in small quantities. It can use small volume cells for the measurement in far-UV and near-UV spectral regions (180-350 nm), which is not possible with bench-top CD.[73] What makes it innovative and unique in comparison with other beamlines worldwide is the installation of a horizontal X-Y motorised stage (Figure 1.18(a)), which can be used for CD *imaging* (CD*i*) at unprecedented 50 μm spatial resolution which is unattainable with bench-top CD instruments (Figure 1.18(b)).[98] The chamber also has control of the temperature in order to be used either in high or cryogenic temperature conditions (-150°C to 300°C) and it can adopt either microtiter or microwell plates to achieve higher work rate screening.[99]

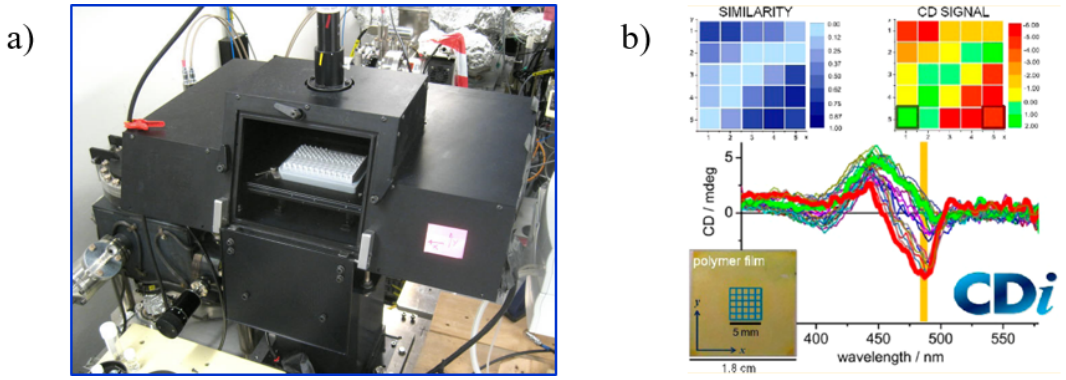

Figure 1.18: a), Unique B23 vertical sample chamber equipped with a motorised XY stage for imaging sample films. **b)**, CD *imaging* (CD*i*) of a polymer film using B23. Reprinted from[99]

1.6 Atomic Force Microscopy, AFM

Atomic Force Microscopy (AFM) is a versatile tool for the characterisation and imaging of surfaces. It was developed by Gerd Binnig, Calvin Quate and Christoph Gerber in 1986 for which they won the Noble Price. Their idea was to create a microscope, which would investigate surfaces on an atomic scale.[100] This kind of microscopes, known as *Scanning Probe Microscopes* (SPMs,) can produce images with atomic resolution. AFM performs through forces which are generated between the surface of a sample and the

probe.[101] Therefore, it is a near field approach and the term *"Atomic Force Microscopy"* (AFM) is justified, since this technique measures the interaction forces when a micro-fabricated sharp tip, referred to as *AFM tip*, probes the surface of the sample into close proximity (μm range) (Figure 1.19).[102]

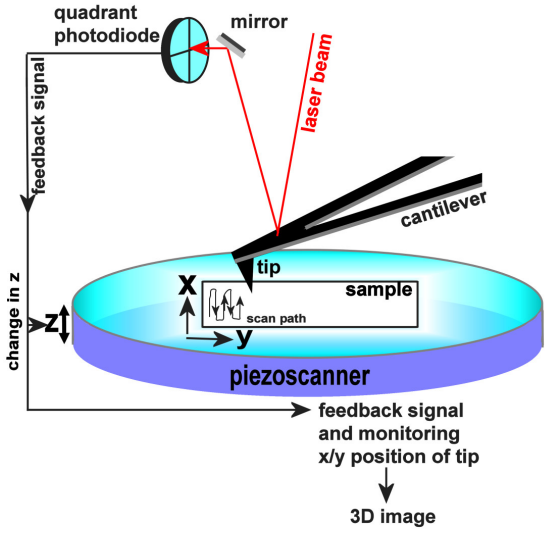

Figure 1.19: AFM schematic. Reprinted from[102]

As mentioned earlier, Seeman first used DNA origami structures in 1998 to image 2D DNA lattices and since then, AFM has become the most popular and reliable tool for the characterisation and imaging of DNA nanostructures.[103] In addition, some other well-known techniques can often be used for visualising structural properties of DNA origami, such as Scanning and Transmission Electron Microscopy (SEM & TEM respectively) and High Resolution Fluorescence Microscopy. The electron microscopies, for example, are chosen for imaging metal containing DNA origami scaffolds because of the high contrast between the biological and metallic materials.[104] However, AFM has the ability to visualize the structures in their native environment and at a nanometre resolution.

One of the biggest challenges while imaging DNA nanostructures using AFM was to overcome the DNA's instability caused by mounting onto the surface, which leads to the microscope tip to exert forces that removed most molecules from the scanned area.[105] Thus, the surface where the DNA origami will be added plays the most important role. An atomically flat, strongly absorbing surface is required and it was found that AFM images were best obtained when DNA was deposited on a mica surface. Using a scotch tape the top layer of mica can be easily cleaved to retain a clean, negatively charged surface.[106] The DNA strands are negatively charged due to the presence of phosphate groups on the DNA backbone. However, the DNA adheres strongly to the mica when a "salt bridge" is created between the negatively charged mica and DNA, which is mediated by the presence of divalent cations (Mg^{2+}, Ni^{2+}, Ca^{2+}) in buffer. The force

is found to be attractive, provided the cations are divalent or higher valency, as high density of monovalent cations has been found to inhibit the attraction of DNA to the mica surface, since a salt bridge cannot form efficiently.[107]

1.7 Aims and objectives

The overall aim of this thesis is to develop a protocol which will allow to orient a protein of interest on a surface, such as DNA nanostructures, and study the complex system with SRCD spectroscopy in order to correlate the SRCD spectral features to specific orientation of proteins. This platform technology will open a door to the study of a number of proteins that require specific orientation for their function, such as membrane proteins. The aims of the research presented in this thesis can be summarised as follows:

1. Adaption of a suitable long range DNA nanostructure which will be utilised as a platform for protein attachment with precise orientation
2. Utilisation of various azF engineered proteins and investigation of immobilisation onto the DNA templates
3. Characterisation of the DNA-protein system using SRCD spectroscopy at B23 Beamline
4. Exploration of the feasibility to create ordered films of DNA nanostructures on surfaces for SRCD studies

1.8 Thesis structure

Chapter 2 investigates the 2D DNA nanogrid as the first DNA template for organising proteins, and more specifically for the attachment of an engineered Green Fluorescent Protein (GFP), along with the first CD data recorded for this system. The B23 CD *imaging* (CD*i*) method is also introduced which allows the probe of different areas of a film sample. Challenges faced with the 2D DNA nanogrid system were solved in Chapter 3 by adapting a different DNA nanostructure, the double-crossover (DX) array and multiple variants of cytochrome b_{562} protein for the hybridisation to the DNA template. CD data of the standalone DX-array or with protein attached are also presented as they were captured both in solution and in films using the CD*i* method. Chapter 4, introduces an alternative approach for immobilisation of the DX-array with and without protein onto the quartz crucibles used for the SRCD experiments based on the fluorous-fluorous interactions. The thesis is concluded in the Chapter 5 which presents the conclusions which can be drawn from this thesis, as well as future work

on the field of bio-nanotechnology and protein studies. Finally, all the experimental procedures carried out throughout this research are outlined in detail in Chapter 6.

Chapter 2

2D DNA Nanogrid for Protein Attachment

The formulation of reliable protocols for the folding of DNA into 2D and 3D structures by exploiting the Watson-Crick base pairing rules led to complex structures, the DNA origami tiles as they are called. The DNA origami tiles are used as a well-defined and addressable template and they have arisen as a versatile tool in bio-nanotechnology for assembling a range of molecules, including proteins.

The hybridisation of the protein on DNA origami tiles is critical, since the attachment of the addressing ssDNA strands to proteins is still limited. Various protein-DNA conjugation methods have been discussed in Section 1.4 of the Introduction, however, a recently explored method is getting more attention. The research of Marth *et al.* has shown that one DNA strand can be attached to a precisely defined residue in a protein, including enzymes, using copper-free click chemistry. This protein modification with the ssDNA allows nanoassembly on base materials such as DNA origami tiles, in order to study the activity of enzymes based on well-defined orientations, tile placement and stoichiometries.[24]

This method of attachment is yet to be applied to extended DNA nanostructures and, subsequently, to structural studies of proteins using SRCD spectroscopy. The idea of this PhD project has evolved from the protein-DNA conjugation method described above, which would be employed in order to orient proteins on more extended and programmable DNA nanostructures. This approach will be transferred to surfaces, where the DNA nanostructures with the attached proteins will be absorbed and, subsequently, studied with SRCD spectroscopy in order to gain a unique insight into the CD response of proteins with fixed positions.

In this chapter, the first DNA nanostructure, which belongs to the family of programmable self-assembling of DNA tiling lattices and which is chosen as a template

for organising proteins, is presented, along with the first CD experiments. As such, the aims addressed in this chapter are:

1. Synthesis of the chosen DNA nanostructure and its basic analysis
2. Modification of a proof-of-concept protein with a ssDNA, attachment of the modified protein to the DNA nanostructure and investigation of the system for successful hybridisation
3. Recording of the first CD data of the DNA nanostructure (unmodified or with protein attached) both in solution and in solid state
4. Determination of suitability of the chosen DNA nanostructure for the purpose of this project

2.1 2D DNA Nanogrid Design Overview

The first DNA nanostructure system which was tested was the 2D DNA nanogrid designed by Hao Yan *et al.*[52] The composition of DNA tiles known as four by four (4 x 4) resembled a cross structure composed of four arm DNA branch junctions. In this project, a two-tile system is used and the tiles are self-assembled into 2D nanogrids. These templates are utilised for protein and enzyme binding. In order to achieve the conjugation of proteins and enzymes to the scaffold, one of the tiles contains an extended strand to allow the attachment, as shown in Figure 2.1). The conjugation will be accomplished when a complementary strand will be introduced to the extended strand. This system was selected for two main reasons. First, the anchoring point, which in our case is used for the protein hybridisation, had already been introduced, hence no complex modification of one of the tiles was required. Finally, only one tile contained the anchoring point and, as a result, there is a larger spacing between the neighbouring proteins, avoiding any undesirable interaction.

2.2 Results and discussion

2.2.1 Synthesis and analysis of the oligonucleotides

A set of 19 DNA single strands, 10 for CTA tile and 9 for CTB tile, were synthesised (Figure 2.2). The CTA-1 strand contained the extended strand, I. The sequence of the extended strand differs from the one that Hao Yan and his group proposed, as it only had 15 A bases, whereas the strand used in this project included A and G bases, for increased stability.

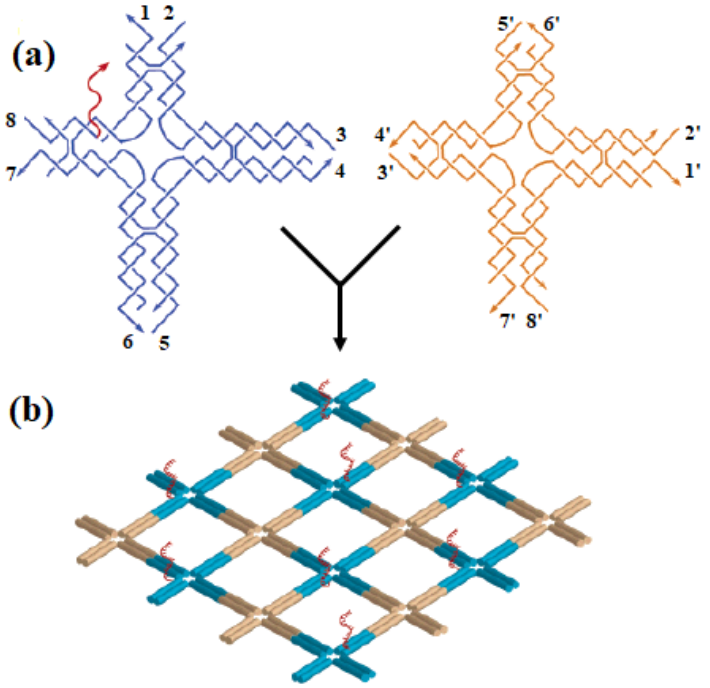

Figure 2.1: **a)** Two tile system to form **b)** the 2D nanogrid. In Subfigure (a), the single strand in red is the extended strand for the conjugation of proteins and enzyme. # 1-8 and # 1'-8' are the designated sticky ends. Reprinted from [52]

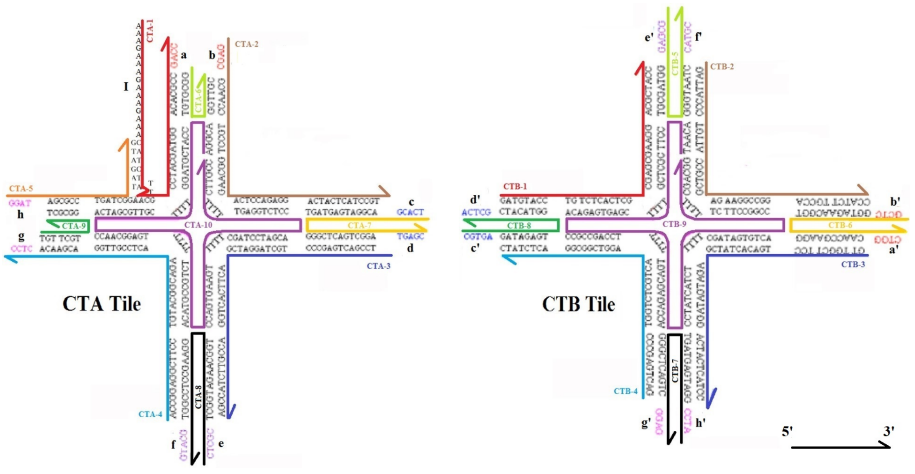

Figure 2.2: **Left**, Schematic of cross tile A (CTA), which consists of 10 strands. **Right**, Schematic of cross tile B (CTB), which consists of 9 strands. Reprinted from. [52]

The sequences of each strand included in CTA and CTB can be found in the Table 2.1.

The DNA synthesis was carried out using an ABI 392 DNA synthesiser and all the strands were synthesised 5' DMT-ON in order to be purified using Glen Pak purification columns. As soon as the protecting groups were cleaved from the bases, Glen Pak purification was performed as indicated in the protocol (see Section 6.3).

Strand	No. of bases	Sequence (5' to 3')	Strand	No. of bases	Sequence (5' to 3')
CTA-1	48	AAA GAA AGA AAG AAA CAT CTA TTA ACG CCT ACG ATG GAC ACG CCG ACC	**CTB-1**	37	GAT GTA CCT GTC TCA CTC GCG AGC GAA GGA CGC TAC C
CTA-2	44	GAG CGC AAC CTG CCT GGC AAG ACT CCA GAG GAC TAC TCA TCC GT	**CTB-2**	40	GAT TAC CCT GTT ACC GTC GAG AAG GCC GGA CCG TTC TACC
CTA-3	47	TCC GAC TGA GCC CTG CTA GGA TCG ACT TCA CTG GAC CGT TCT ACC GA	**CTB-3**	42	GTT GGC TTC CTG ACA CTA TCG AGA TGA TAG GAC TAC TCA TCC
CTA-4	45	ACC GGA GGC TTC CTG TAC GGC AGA ACT CCG TTG GAC GAA CAC TCC	**CTB-4**	39	GAC TGA GCC CTG GTC TCG TCA AGG TCG GCG GAC TCT ATC
CTA-5	23	GGA TAG CGC CTG ATC GGT AGA TG	**CTB-5**	26	GCG AGG GTA GCG TGG GTA ATC CAT GC
CTA-6	13	GGC GTG TGG TTG C	**CTB-6**	29	GCT CGG TAG AAC GGT GGA AGC CAA CGG TC
CTA-7	36	TCA CGA CGG ATG AGT AGT GGG CTC AGT CGG ATG AGC	**CTB-7**	29	ATC CGG ATG AGT AGT GGG CTC AGT CGG AG
CTA-8	36	CTC GCT CGG TAG AAC GGT GGA AGC CTC CGG TGC ATG	**CTB-8**	26	CGT GAG ATA GAG TGG TAC ATC GCT CA
CTA-9	13	TGT TCG TGG CGC T	**CTB-9**	100	TAA CAC CTT CGC TCG TTT TCG AGT GAG ACA CCG CCG ACC TTT TTT GAC GAG ACC ACC TAT CAT CTT TTT CGA TAG TGT CAC CGG CCT TCT TTT TCG ACG G
CTA-10	100	AGG CAC CAT CGT AGG TTT TCG TTG CGA TCA CCA ACG GAG TTT TTT CTG CCG TAC ACC AGT GAA GTT TTT CGA TCC TAG CAC CTC TGG AGT TTT TCT TGC C			

Table 2.1: The first 3 columns from the left-hand side present the sequences of the oligonucleotides for the CTA tile, whereas the 3 columns from the right-hand side present the sequences of the oligonucleotides for the CTB tile.

Following the purification, UV-Vis spectroscopy was used to measure the absorbance of a DNA sample. Given the absorbance, the concentration of the oligonucleotides was then calculated using the Beer-Lambert Law (eq. 2.1), and therefore, the number of moles synthesised was obtained.

$$A = \varepsilon cl \tag{2.1}$$

where A is the absorbance, ε is the extinction coefficient ($M^{-1}cm^{-1}$), c is the concentration in (M) and l is the path length (cm).

The path length is a known value (1 cm) and the extinction coefficient is determined using the ATDbio oligo calculator. In Figure 2.3, the UV-Vis spectra of CTA and CTB are shown. The peaks observed at 260 nm correspond to the DNA.

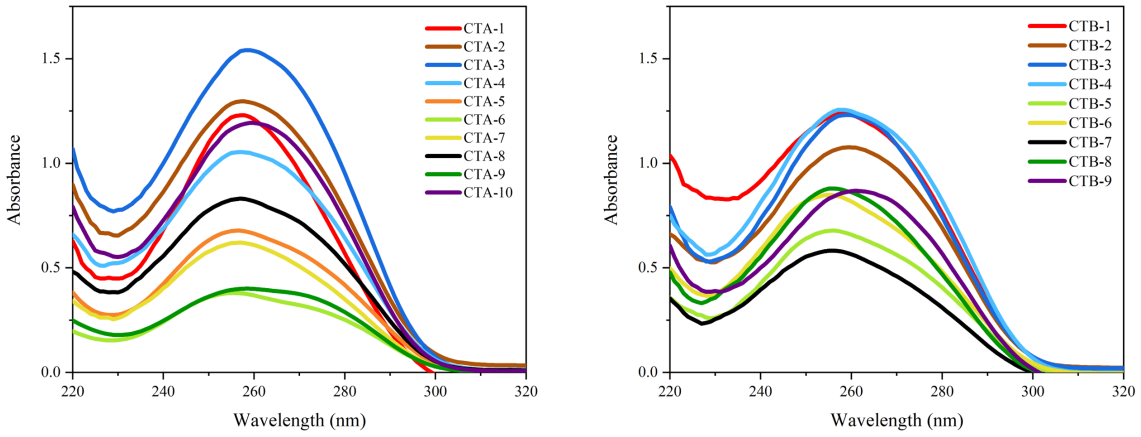

Figure 2.3: UV-Vis spectra of unmodified DNA with a peak at 260 nm for DNA. UV-Vis spectra of the strands involved in CTA tile **(left)** and in CTB tile **(right)**.

Once the concentration is calculated and since the volume is known, the number of moles of DNA can be obtained using the following equation.

$$Number\ of\ moles = volume \times concentration \tag{2.2}$$

The DNA samples were further analysed using mass spectrometry (MS) which was carried out by the MS group of the Chemistry Department at the University of Southampton. The masses were calculated using the Oligo calculator from ATDBio.[108] The mass of each strand found using mass spectrometry is shown in Table 2.2. Traces from all DNA strands included in CT-A and CT-B tiles can be found in the Appendix A. The table also shows the number of moles of DNA and the synthesis yield.

2.2.2 Formation of 2D DNA cross-tiles and nanogrid

After the synthesis and analysis of the DNA strands, their assembly into the 2D DNA nanogrid could be performed. Different protocols of assembly in solution were tested.

CTA Tile	Calculated Mass (g/mol)	Mass Found (g/mol)	Number of moles (nmol)	Yield (%)
CTA-1	14773.6	14774.3	459.0	46
CTA-2	13471.6	13471.2	579.6	58
CTA-3	14351.2	14351.1	483.6	48
CTA-4	13791.8	13792.0	452.4	45
CTA-5	7144.6	7143.8	535,8	54
CTA-6	4037.6	4036.4	664.8	67
CTA-7	11206.2	11205.1	315.6	32
CTA-8	11118.1	11118.1	447.6	45
CTA-9	3972.6	3972.4	638.4	64
CTA-10	30679.6	30679.5	234.0	23
CTB Tile	**Calculated Mass (g/mol)**	**Mass Found (g/mol)**	**Number of moles (nmol)**	**Yield (%)**
CTB-1	11335.3	11335.1	643.2	64
CTB-2	12232.8	12231.6	522.6	52
CTB-3	12864.3	12863.3	558.0	56
CTB-4	11975.7	11974.3	634.2	63
CTB-5	8092,2	8091.6	484.2	48
CTB-6	8992.8	8994.6	539.4	54
CTB-7	9038.8	9037.6	364.2	36
CTB-8	8035.2	8034.5	607.8	61
CTB-9	80544.5	30544.2	174.0	17

Table 2.2: Analytical data of Mass Spectrometry for CTA and CTB.

According to the Protocol F1 (Section 6.10), all the single strands involved in the tiles were mixed in an Eppendorf tube with 1 x TAE/ Mg^{2+} buffer in total volume of 60 μL and concentration 1 μM. The mixture was heated up to 90 °C and cooled down to 4 °C in a period of ∼13 hours on T100 Thermal Cycler.

Based on Protocol F2 (Section 6.10), each individual tile (CTA & CTB) was assembled by mixing a stoichiometric quantity of the DNA strands was in 1 x TAE/Mg^{2+} buffer. The oligo mixtures were cooled from 90 °C to 4 °C in a period of ∼13h using the T100 Thermal Cycler (similar Protocol F1). Then, the tiles were mixed with a ratio of 1:1, heated up to 50 °C and cooled to 4 °C in a period of 6 hours, in order to form the aforementioned 2D DNA nanogrid.

Agarose gel electrophoresis (1.0%) was performed to visualise the fragments and give an insight into whether the individual tiles and the extended grid were formed. The gel showed that the self-assembly was completed to a certain extent. According to the gel image (Figure 2.4), the individual CTA and CTB tiles were obtained. The 2D nanogrid was also formed, regardless of the protocol used, as shown in the fragment at the top of the gel. However, it did not run through the gel due to its large structure compared to the single tiles. Finally, the individual tiles and strands can also be observed, moving through the gel.

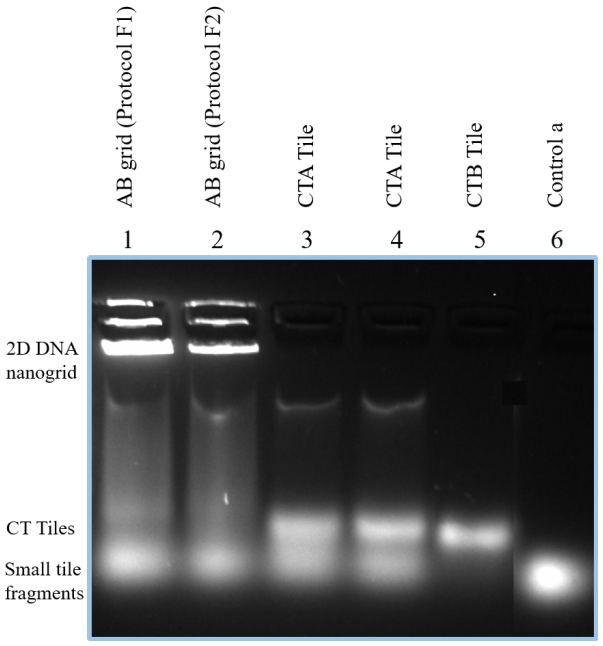

Figure 2.4: 1.0% agarose gel. Lane 1: AB grid annealed with Protocol F1, Lane 2: AB grid annealed with Protocol F2, Lane 3: CTA Tile, Lane 4: CTA Tile, Lane 5: CTB, Lane 6: Control a consists of 4 strands involved in tile CTA, but not heated.

2.2.3 Synthesis of the modified oligonucleotide

In order to attach the protein on the DNA nanogrid, the protein must be modified with a ssDNA, which is complementary to the extended strand I involved in CTA tile. For this purpose, the ssDNA must be first modified with an alkyne, which in turn will react with the protein via Click Chemistry. The modifier used in this research is the 5'-Click-easy™ BCN CEP II (Figure 2.5). This modifier can be used to "prepare" the oligonucleotides for the strain-promoted alkyne-azide cycloaddition (SPAAC).

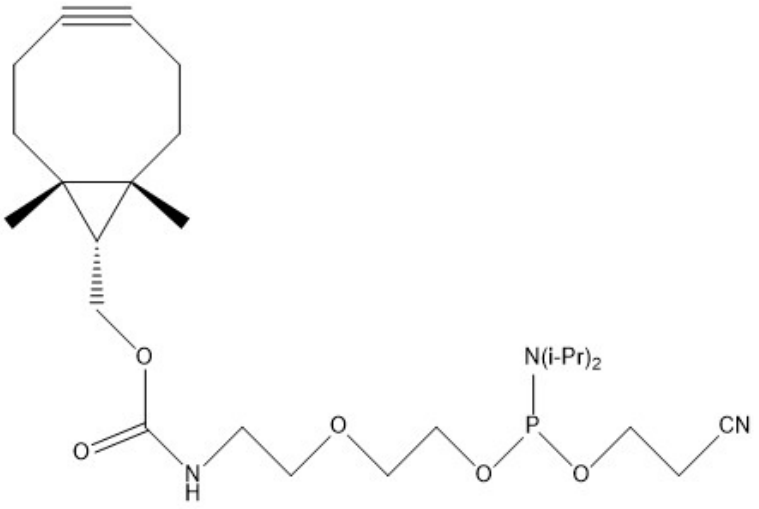

Figure 2.5: Strained bicyclononyne (BCN).

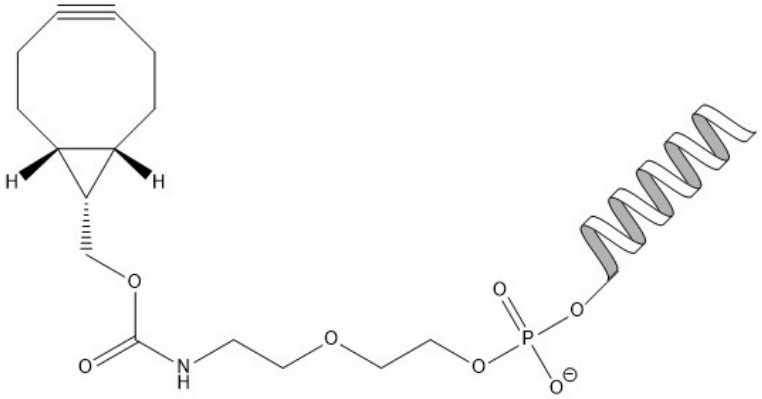

Figure 2.6: The ssDNA with the terminal bicyclononyne (BCN) for subsequent copper-free click elaboration.

The synthesis of the DNA strand and its modification were carried out on an Applied Biosystems Expedite DNA synthesiser. The modifier, 5'-Click-easy™ BCN CEP II, was dissolved in anhydrous acetonitrile, vortexed and then, it was added to the DNA synthesiser. Extended coupling time up to 8 minutes was used for the modification of the oligonucleotide. The modifier was added at the 5'-end of the desired strand. The modified oligonucleotide was synthesised twice in order to have more material available for the experiments. The sequence of the modified DNA is presented in Table 2.3.

Strand	No. of Bases	Sequence (5 '- 3')
5' BCN-DNA 1	18	YTTTTTTCTTTCTTTCTTT
5' BCN-DNA 2	18	YTTTTTTCTTTCTTTCTTT

Table 2.3: Sequence of the modified oligonucleotide, where Y is the 5'-Click-easy™ BCN CEP II .

The BCN-DNAs were cleaved from the beads using aqueous ammonia and they were then left for deprotection in the thermomixer. The deprotected modified oligo was evaporated to dryness using the speed-vac and, subsequently, desalting procedure was performed using Glen-Pak (see protocol in Section 6.3) to remove excess salt. The ssDNA with the modification is shown in Figure 2.6.

Following the desalting procedure, the absorbance of the BCN-DNAs was measured using UV-Vis spectroscopy. The concentration of the oligonucleotides was calculated using the Beer-Lambert Law, as mentioned in Section 2.2.1.

In Figure 2.7, the UV-Vis spectra of the BCN-DNAs are shown. The peaks observed at 260 nm correspond to the DNA.

The number of moles of DNA was also obtained and the modified oligonucleotides were further analysed using mass spectrometry. For the calculation of the masses the ATDBio oligo calculator was used here as well. The mass of each strand found using mass

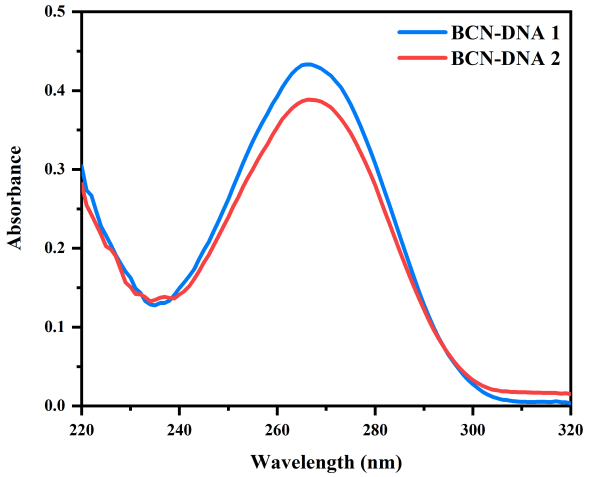

Figure 2.7: UV-Vis spectra of BCN - DNA with a peak at 265 nm.

spectrometry is shown in the following table (Table 2.4). Traces from the MS analysis can be found in Appendix A, Figure 20.This table also shows the number of moles of DNA and the synthesis yield.

	Calculated Mass (g/mol)	Mass Found (g/mol)	No of moles (nmol)	Yield (%)
5' BCN-DNA 1	5710.8	5711	197.2	19.7
5' BCN-DNA 2	5710.8	5711	102.4	10.2

Table 2.4: Analytical data of Mass Spectrometry of the BCN-DNA.

It is important to highlighth a mistake that was made in the synthesis of the modified DNA, which was used for the click chemistry with the protein. As mentioned in subsection 2.2.1, the extended strand included on the CT-A tile consists of 15 bases and it would be sensible that its complementary strand (BCN-DNA) also consists of 15 bases. However, an 18-base long modified strand was synthesised, including 3 more thymine nucleobases at the 5 end, which was a mistake only noticed at a later stage, but it did not affect the hybridisation of the protein on the tile. This mistake was corrected when the strands were synthesised again for use with the new DNA nanostructure in Chapter 3.

2.2.4 Protein Modification- Click Chemistry

As soon as the BCN-DNA was synthesised and analysed, it was used for the modification of the protein, Green Fluorescent Protein (GFP). GFP is an easy-to-handle protein that is easily detectable by fluorescence spectroscopy and tolerant to high temperature. A modified variant of GFP residues, the superfolder GFP (sfGFP, MW $\sim$ 27kDa) was modified with p-azidophenylalanine (azF) at position 204 (Figure 2.8) using

an expanded genetic code system by our collaborators, Dafydd Jones and his group from the University of Cardiff.[69,109]

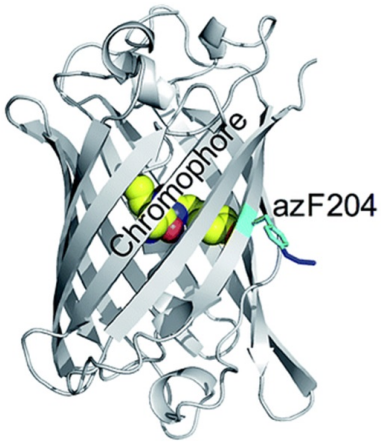

Figure 2.8: Structure of sfGFP^{azF204}. Residues critical to function are shown as spheres, and residue targeted for replacement with azF is shown in stick representation. Adapted from[110]

In order to introduce the BCN-DNA at the defined site of the GFP, a bio-orthogonal and biocompatible copper-free strained ring promoted alkyne-azide cycloaddition (SPAAC) reaction was carried out (Figure 2.9).

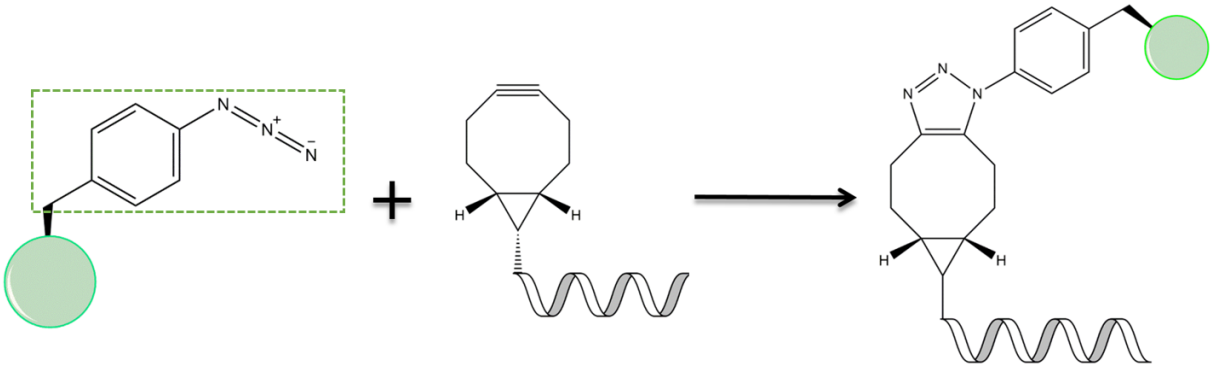

Figure 2.9: The protein contains a genetically encoded p-azido-L-phenylalanine (azF; enclosed in box) and the ssDNA contains the BCN. The two components create a triazole link.

SfGFP^{azF204} was mixed with modified 5'-BCN for 48 hours at room temperature in the dark. The sfGFP variant was successfully modified with the BCN-DNA. The progress of the modification was monitored using sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) (Figure 2.10). The percentage refers to the amount of protein or protein-DNA sample included in 10 μL total volume of loading buffer and protein or protein-DNA which was loaded on the gel.

The mixture was concentrated using Amicon filters. Following, the absorbance and the fluorescence intensity were measured using UV-Vis and Fluorescence spectroscopy, respectively. The spectra are presented in Figure 2.11. For the UV-Vis spectra, the

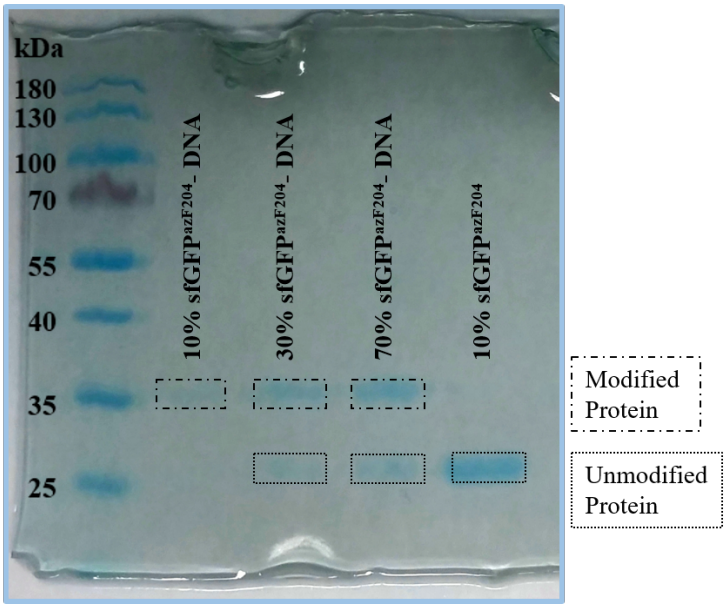

Figure 2.10: Gel mobility shift analysis of protein-DNA conjugation. The left hand lane is the molecular weight standards. Lanes represented the unmodified protein and the protein-DNA are annotated on the figure. Three different concentrations were used for the protein-DNA. The MW of the sfGFP^{azF204} and the sfGFP^{azF204}-DNA is ∼27.7 kDa and ∼37 kDa, respectively.

peaks at 265 nm and 495 nm correlate to the DNA and the GFP, respectively. The shape of the intensity signal is similar to the ones found in Marth's work,[24] which has also shown that the DNA has very little impact on the function of the protein, since the overall fluorescence intensity of unmodified and modified GFP was very similar.

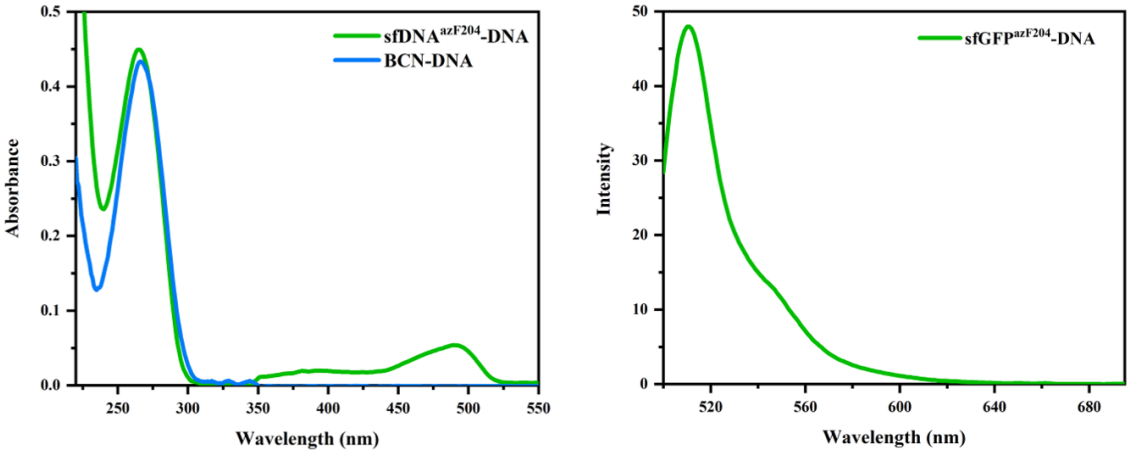

Figure 2.11: Left, UV-Vis spectrum of the protein-DNA with a peak at 265 nm for the DNA and at 495 nm for the sfGFP^{azF204} and UV-Vis spectrum of BCN-DNA with a peak at 265 nm, **Right,** Fluorescence of protein DNA, emission spectrum was recorded on excitation at 485 nm.

2.2.5 Protein-ssDNA co-assembly on DNA nanogrid

The 2D DNA nanogrid was formed using F2 protocol described in section 6.10 and it was mixed with sfGFP conjugated with addressing ssDNA, complementary to the extended strand I of the CTA tile. 70% of sfGFP^{azF204} was used for the first experiment. The solution was annealed either from 37 °C to 4 °C on Thermal Cycler, or at room temperature.

The system was then visualised using Atomic Force Microscopy and it was further analysed with Synchrotron Radiation Circular Dichroism spectroscopy.

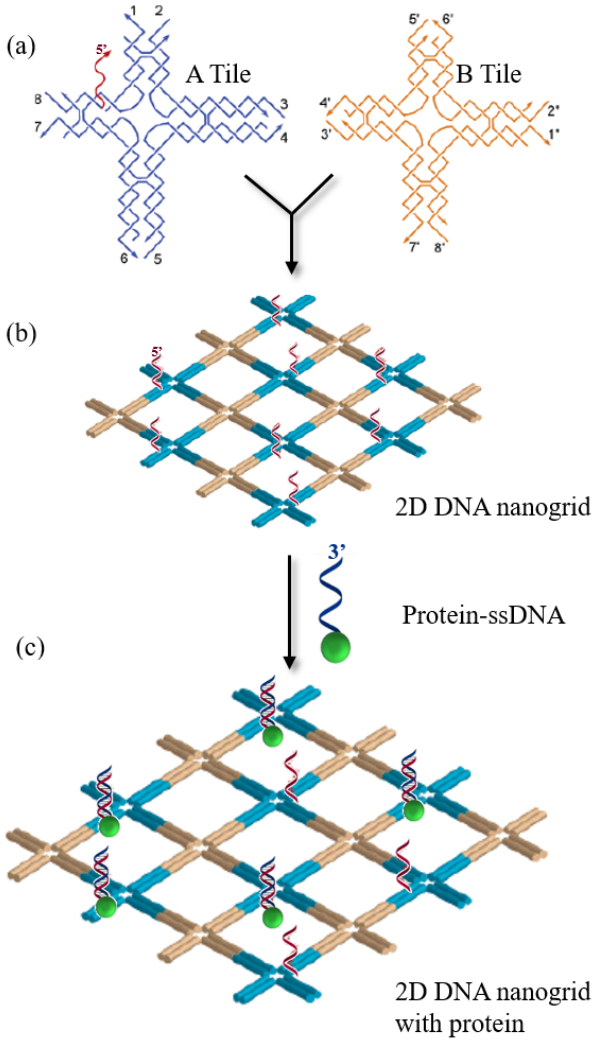

Figure 2.12: **(a)** The two tile system to form the 2D DNA nanogrids. **(b)** The 2D DNA nanogrid with the extended strand on each A tile pointing out of the plane. **(c)** Assembly of the protein on the DNA grids. Adapted from [52]

2.2.6 Atomic Force Microscopy of 2D DNA nanogrid

The accomplishment of the self-assembly were assessed with Atomic Force Microscopy (AFM). The main goal was not only to image the DNA nanostructure, but also to determine the optimal conditions to produce the best quality of samples and images. For the experiments that were performed for this project, all images were obtained using tapping mode AFM, both in air and in liquid. Tapping mode was developed to image soft biological samples at high resolution. It reduces the lateral forces that are exerted on the sample since the tip spends less time on the surface of the sample in contrast to the contact mode that is constantly in contact with it (Figure 2.13) The cantilever vibrates at its resonance frequency, so it effectively taps as it travels over the sample resulting in a decrease of its oscillation amplitude.[111] This is detected by the photodiode and the image of the sample surface is produced. As the tip is not in constant contact with the sample, the chances of deformations are minimised as it exerts less force on the sample and can image loosely bound objects on the substrate.[112] Consequently, this mode is preferred to contact mode when imaging samples as DNA nanostructures.

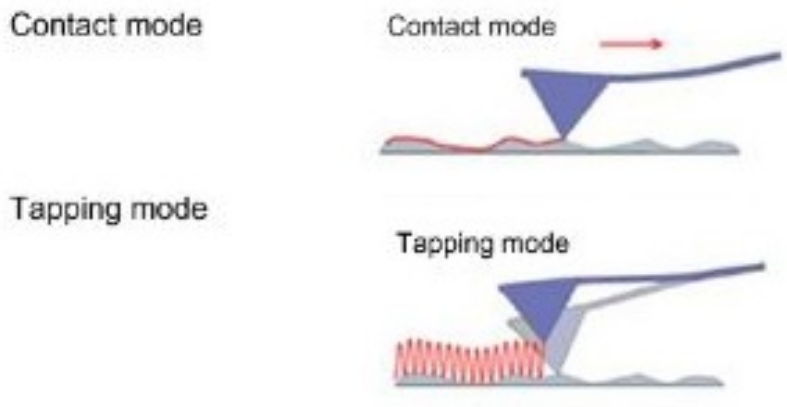

Figure 2.13: Different modes of AFM, including contact and tapping mode. Adapted from[113]

According to Yan *et al.*, the distance between CTA and CTB tiles is supposed to be 4 or 4.5 helical turns of DNA plus the diameters of the DNA between the adjacent tile centres, totalling approximately 17.6 nm or 19.3 nm.[52] The 2D DNA nanogrids that were self-assembled in solution were imaged by being deposited on a cleaved mica surface, which is a negatively charged material and commonly used to immobilise DNA via a charge bi-layer. The buffer that was used for the absorption on the mica surface was 1X TAE/ Mg^{2+}. The magnesium ions present in the buffer create a positive layer on mica, allowing the DNA nanostructure to be immobilised on the surface.[114,115] Different areas of the mica surfaces with the deposited sample were scanned randomly in order to collect as many data as possible and capture the most representative images of the DNA nanostructures with and without the protein attached.

2.2.6.1 AFM under dry conditions

The visualisation of the nanogrid with and without the protein attachment was first obtained with AFM under dry conditions.

The 2D DNA nanogrid (1 μM) was formed using the Protocol F2 (Section 6.10) and subsequently, it was imaged with the Nanoscope III SPM controller AFM at the University of Southampton. 10 μL of 1 μM DNA nanostructure solution was deposited on a freshly cleaved mica surface and dried using a nitrogen gun. The images captured are presented in Figure 2.14(a) and 2.14(b). Some cross structures can be observed in random parts of the mica surface, but they are not similar to the nanogrid reported in literature (Figure 2.14(c)). The major part of the sample area seems disordered with high nanostructure density and some areas resembling the desirable nanogrid structure which has either not fully formed or was damaged by the AFM tip.

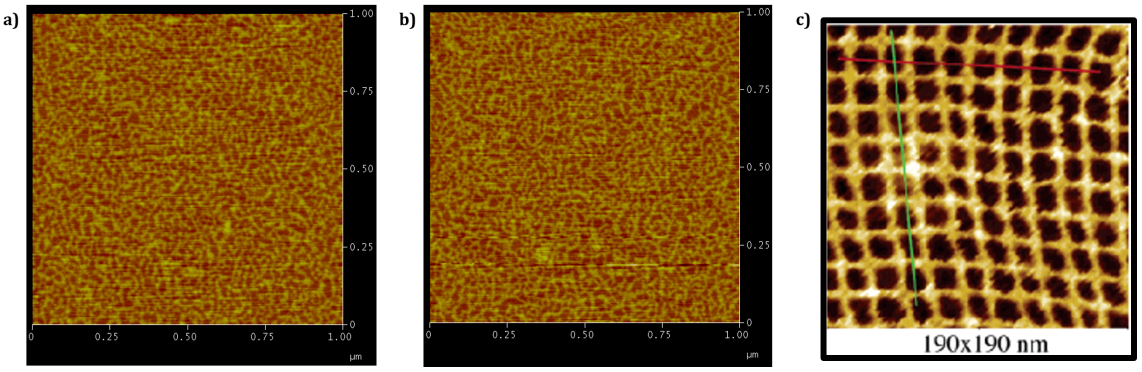

Figure 2.14: a),b) AFM images of 2D DNA nanogrid (1 μM)as captured using AFM in air (1 x 1 μm). **c)** AFM image of 2D DNA nanogrid as captured by Yan *et al.* (190 x 190 nm).[52]

The DNA nanogrid with the protein attached was imaged in air using FastScan Bio™ AFM from Bruker at Strathclyde University in Glasgow. The nanogrid was formed using the Protocol F2 and then it was co-assembled with the protein-ssDNA (70%), as described in Section 6.12. This time, 5 μL of sample was disposed on the mica, in order to check if the density of the sample on the mica would change. In Figure 2.15(A), there are a few visible cross structures (enclosed in squares). However, the sample was too concentrated and no periodicity could be observed. Hence, it is not clear whether the nanogrid was successfully formed.

Following, the ochre and white spots, which are visible in both images, might be either the attached protein or buffer salt that remained even after washing the mica surface with deionised water. In Figure 2.15(B), the proteins are more detectable than the nanogrid which can help to investigate if the distance between neighbouring proteins is similar to literature. The zoomed-in area of Figure 2.15(B) presents 3 protein dots whose spacing is measured according to the section analysis at ~38 nm in linear direction (blue-to-blue tile, Figure 2.12(c)) which is in accordance with the spacing of neighbouring

nanoparticles in Yan's study.[52] The height profile, which is 1.5 nm, can also indicate the presence of protein since the nanogrid should have a height of 1.0 nm. The difference in the height after the hybridisation of the protein may not be very large, but it could be because the protein is located closer to the surface of the tile and subsequently it is facing towards the mica surface. Nevertheless, since the sample was highly concentrated, it was hard to locate another series of proteins in line for comparison with the one described above and therefore, it cannot be safely concluded that all the proteins observed in the figure were bonded to the tile and the three dots were not accidentally in line.

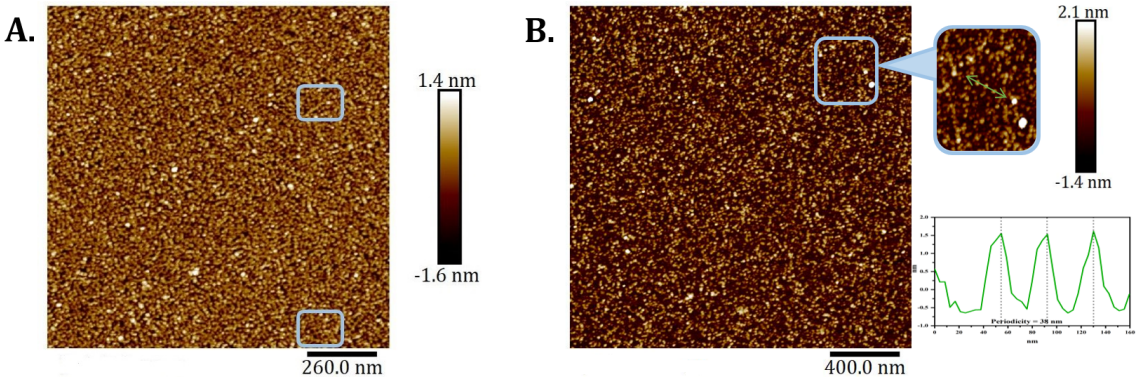

Figure 2.15: Images of 1 μM DNA nanogrid with protein attached **A.** Image 1.3 x 1.3 μm. Cross structures are visible and eclosed in blue squares. **B.** Image 2.0 x 2.0 μm. The protein dots are more visible here than the nanogrid. The section analysis presented is for the zoom-in area that shows 3 protein dots with a distance of 38 nm from each other.

2.2.6.2 Liquid AFM

Liquid AFM was used to obtain useful information regarding the structure of the 2D DNA nanogrid, as it is able to offer better resolution compared to the dry AFM.

The sample of the 2D DNA nanogrid without the protein was also imaged with tapping mode with liquid AFM (FastScan Bio™ with ScanAsyst from Bruker) at Strathclyde University in Glasgow. 5 μL of 1 μM sample was spotted on the mica and then, 20 μL of the annealing buffer was added on top. Some small cross structures can be observed in random parts of the mica surface, but with no periodicity as presented in Figure 2.16. Important details might have been lost as single strands or tiles might be covering the nanostructures on the surface. Furthermore, DNA assemblies are fragile and, therefore, they might deform or break in smaller pieces during the liquid handling. However, the the higher resolution images of the liquid AFM showed a more visible structure.

The images were further analysed with NanoScope Analysis 1.5 from Bruker in order to obtain structural information such as the distance between the cross tiles.

Figure 2.17 presents the AFM image from Figure 2.16(b). Despite the high density of the sample in the mica surface, some square lattices are clearly visible and especially those

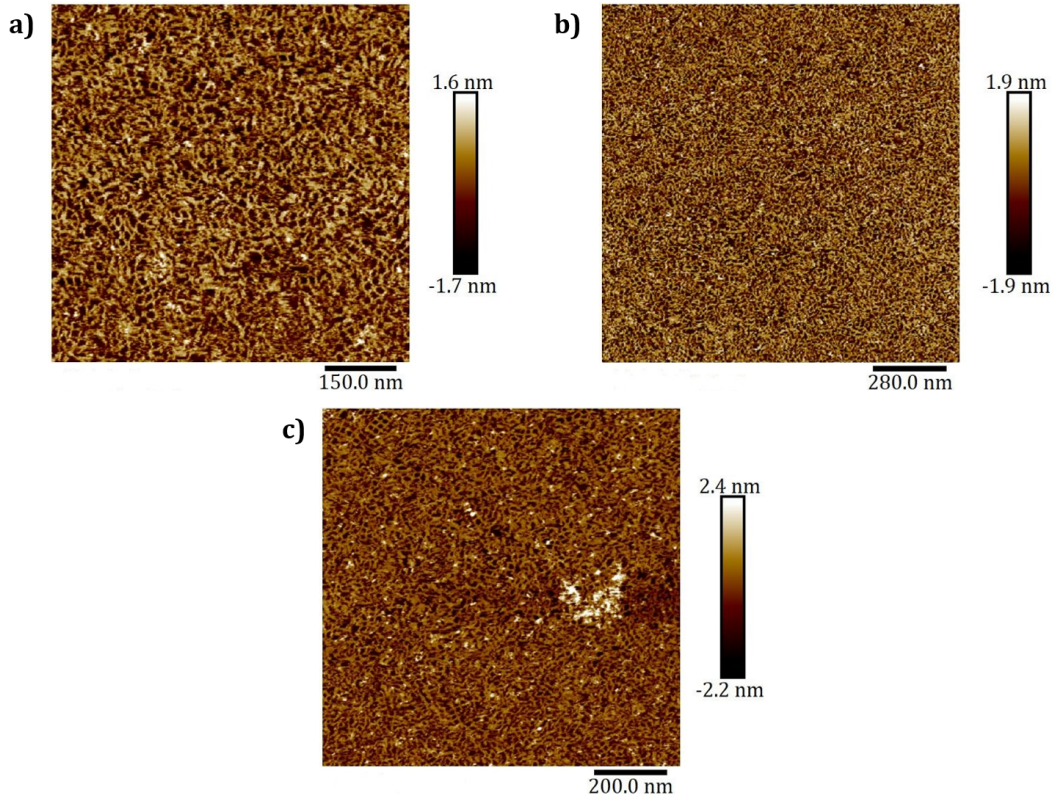

Figure 2.16: Images of 1 μM 2D DNA nanogrid as captured in liquid AFM. **a)** Image 791 x 791 nm. **b)** Image 1.4 x 1.4 μm. **c)** Image 1.0 x 1.0 μm.

which are enclosed in light blue squares. For each zoomed-in area, the related analysis sections are also presented. Area A was analysed for the blue-to-blue tile diagonal repeat (as indicated in Figure 2.17) noted with the yellow arrow. According to the analysis, the periodicity is 25 ($\pm$ 2) nm. In the meantime, area B was analysed for the linear repeat along the red and the blue arrows. Furthermore, the cavities from one centre to the other measure 17 ($\pm$ 3) nm, and subsequently the distance between blue-to-blue tile linear repeat is 34 ($\pm$ 3) nm. All the distances are in agreement with the literature, as shown in Figure 2.17(C) with the DNA nanostructure reprinted from Yan's work.[52]

2.2.7 SRCD & CD Spectroscopy for the 2D DNA crosstiles & nanogrid

The SRCD measurements described below were recorded at B23, Diamond Light Source, UK, during beamtime allocations including SM15269 and SM18017. Measurements were performed both in solution and solid phase (films) using the module B end station. On the other hand, conventional CD spectra were acquired on a nitrogen-flushed Chirascan Plus spectropolarimeter. Both the SRCD and CD data were processed in CDApps[116] and Origin (OriginLab Corporation). More details are presented in Chapter 6 (Section 6.14).

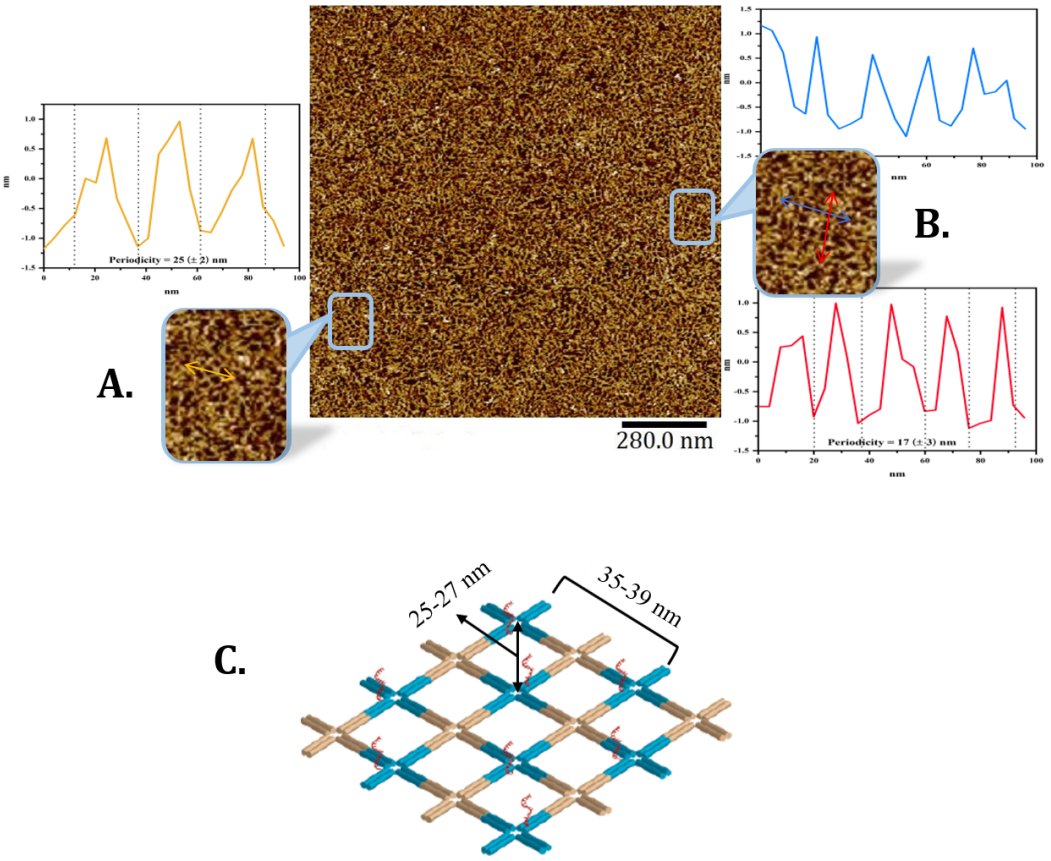

Figure 2.17: Image of 1 μM 2D DNA nanogrid (1.4 x 1.4 μm) from Figure2.16(b) as captured in liquid AFM and the analysis sections for each zoomed area **A.** A zoom-in area showing a square lattice. The section analysis above represents the diagonal distance of the tiles along the yellow arrow. **B.** Another zoom-in area showing a bigger square lattice. The analysis sections below and above describes the linear distance of the tiles along the red and blue arrows, respectively. **C.** The expected distances of the cross tiles (blue-to-blue tile, diagonal repeat: 25-27 nm and linear repeat: 35-39 nm). DNA nanogrid scheme adapted from ref.[52]

In the following subsections, the CD and SRCD spectra of the 2D DNA nanogrid, which is referred to as *AB grid*, with and without GFP attached, as well as the individual crosstiles CTA and CTB, and the sfGFP^{azF204}, are presented as they were measured either in solution or in solid phase.

2.2.7.1 SRCD measurements in solution

A series of preliminary measurements were performed in order to investigate the SRCD spectral features of the individual crosstiles CTA and CTB, and of the AB grid. Two different samples of the AB grid were prepared using the two different formation protocols, Protocol F1 and F2 (as described in section 6.10), in order to compare them and confirm their similarity. Protocol F1 was followed for the formation of the individual tiles.

30 μL (1 μM) of each crosstile (A and B) and of each AB grid was pipetted in a Suprasil quartz cylindrical cell (Hellma) of 0.02 cm pathlength and subsequently, they were measured in solution using the Module B at B23 Beamline at Diamond Light Source. The average of four scans recorded for each sample is presented in Figure 2.18. All the measurements were baseline-corrected with a spectrum obtained from the 1 x TAE/Mg^{+2} buffer.

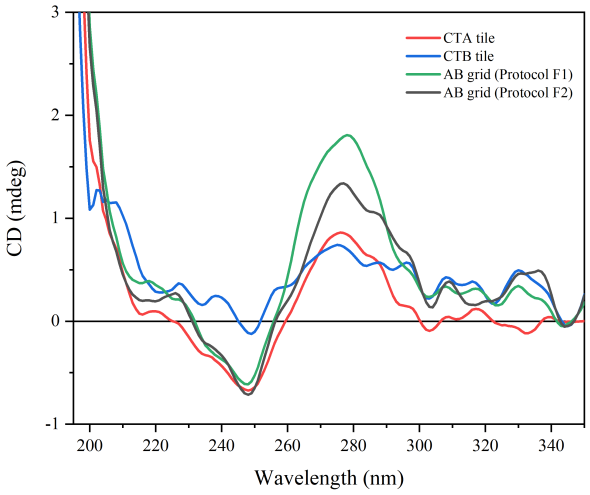

Figure 2.18: SRCD spectra of the CTA (red) and CTB (blue) tiles and the AB grids Protocol F1 & F2 (green and blue, respectively) in solution. Conditions: 1 μM in 1x TAE/Mg^{2+} buffer.

The B-form duplex and consequently the DNA nanostructures that belong in this category are characterised with a CD signal which consists of broad maximum at 260-280 nm and a minimum at 245 nm. The traces of CTA and CTB crosstiles behave in a similar manner. The signatures of the AB grids also behave similarly, with only the intensity differing. The AB grid formed with Protocol F1 has higher intensity at ~275 nm than the one formed with Protocol F2, even though they have the same concentration. Qualitatively, they have the same B form folding but quantitatively it seems that they do not. This may be due to the fact the AB grid protocol F2 has smaller fragments. However, the SRCD signatures recorded for each sample are comparable to the characteristic ones of the DNA with a positive peak at ~275 nm and a negative one at ~245 nm.

It is crucial to record the absorbance on every CD and SRCD signature measurement, in order to ensure that the CD and SRCD spectra are measured above the noise. Therefore, the data points are valid and the optimal conditions for the film preparation (thickness, concentration) can be determined. Nevertheless, the absorbance was not measured in this case, since the purpose of this preliminary experiment was to get familiarised with the behaviour of the system. Yet, Figure 2.19 shows the High Tension (HT) on the photomultiplier detectors of Module B, which can also give information about the reliability of the measurements. When the HT voltage is greater than 600 V, the signal start to become noisy. Greater than 700 V, the noise becomes disproportionate to the signal and the PMT data points are not reliable and should not be used. Ideally, the

HT must be smaller than 600 V. In the case of the samples presented in Figure 2.19, all the HT data from 190-350 nm are below 600 V, hence the SRCD data (Figure 2.18) at this wavelength are reliable and useful. It is important to highlight though that the HT voltage reaches 1000 V below 195 nm and, therefore, there is a cut-off for the SRCD data at 195 nm.

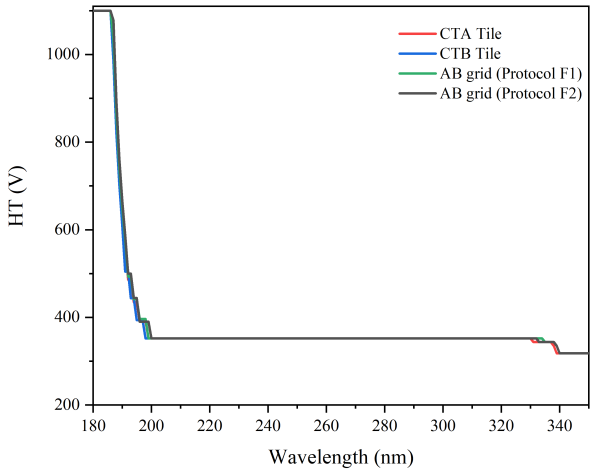

Figure 2.19: High tension (HT) on the photomultiplier detectors of Module B of the CTA (red) and CTB (blue) tiles and the AB grids Protocol F1 & F2 (green and blue, respectively) in solution.

2.2.7.2 Benchtop CD measurements in solid phase

The 2D DNA nanogrids were further analysed in solid phase (films). For the preparation of the films, a fused silica substrate (quartz disk, Supracil) with 22 mm diameter and 1.25 mm thickness was used. This material is widely found in CD cuvettes and substrates due to its inherent strength and excellent transmittance in the UV region, as well as in the preparation of oriented CD samples according to literature.[117] The films of the AB grid (Protocol F1) and the AB grid (Protocol F2) consisted of 3 and 4 layers, respectively, of 5 μL sample. After the addition of each layer, the sample was dried under a stream of nitrogen.

The samples were first tested with the Chirascan Plus (bench top instrument). Two recordings were made on each film, with the second one upon rotating the sample by 90° clockwise. The absorbance was also recorded and the respective results, including the CD data, are shown in Figure 2.20.

The CD spectra of the films showed more complex spectral features rather than a characteristic DNA signature in the B form as shown in Figure 2.18. The characteristic positive peak at ~280 nm is visible, however the negative peak at ~245 nm is almost missing, since its shape is broader and not distinguishable. A positive and a negative peak at 220 nm and 210 nm, respectively, are also visible, which also usually appeared in DNA CD spectra. The intensity among the samples varies and the original samples

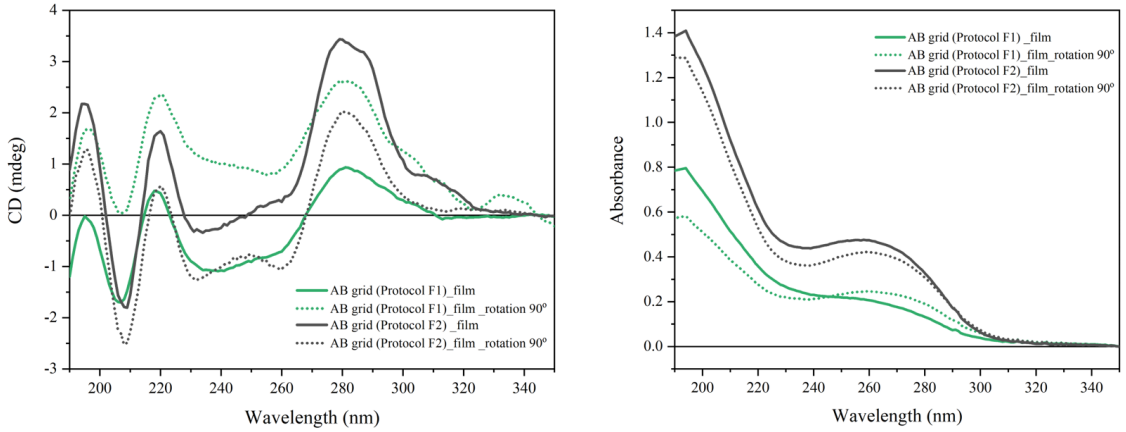

Figure 2.20: Left, CD spectra of AB grid films before and after 90˚ rotation (solid and dotted lines, respectively). **Right**, Absorbance spectra of AB grid films before and after 90˚ rotation (solid and dotted lines, respectively).

lead into different signatures compared to the rotated ones. However, if the CD spectra are offset at around 310 to 320 nm, as shown in Figure 2.21, the solid and dotted grey line for the AB grid formed with Protocol F2 match, and they have the same profile, in contrast to the green ones for the AB grid Protocol F1. Hence, the question which arises is whether Protocol F2 is better than the F1, which will be further investigated in the following sections.

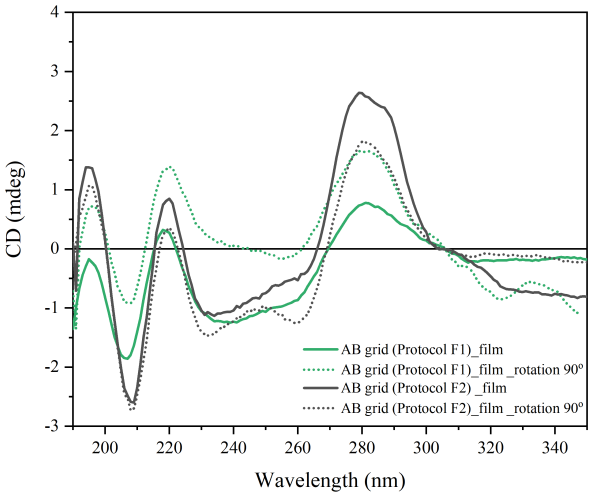

Figure 2.21: CD spectra of AB grid films before and after 90˚ rotation (solid and dotted lines, respectively) after offset at around 310 to 320 nm.

On the other hand, the variations in intensity might be due to different concentrations of the area recorded, which is noticeable from the absorbance spectra (Figure 2.20(right)). The same sample film, but with different film orientation, has dissimilar absorbance which led to the conclusion that the film was not uniform in terms of concentration.

Another important aspect for the CD experiments of the oriented samples when they are absorbed on surfaces as films is that the samples must have good optical properties

in order to avoid the effects of artefacts, which are mostly related to linear dichroism (LD). LD occurs when the fused silica used for the oriented CD studies is tilted and not exactly perpendicular to the incident light source. According to literature, poorly prepared samples with uneven and inhomogeneous layers lead to increased LD, whereas the well prepared samples which should have the DNA helix axis parallel to the electric field, have an LD equal to 0.[118] In the case of the spectra presented in Figure 2.20 and the corresponding off-set data presented in Figure 2.21, the changes of CD due to rotations are indicative of LD contributions that can certainly distort the overall CD spectrum.

These LD artefacts can be minimised by rotating the sample using eight sets of spectra that are collected at every 45°. That kind of experiment was performed using the film of the AB grid (Protocol F1) as it was the one that showed differences in the signature upon 90° rotation from the original. The resulting spectra are presented in Figure 2.22 and each spectrum derives from the average of 5 scans. Usually, for the experiments in solid phase, the blank fused silica substrate is being used as a baseline, but unfortunately those data were missing in this experiment. Thus, the data for the same blank silica substrate but from a different experiment were used, which were also recorded using the Chirascan, in order to subtract them from the samples' data.

However, it is evident from Figure 2.22 that the CD signal is quite different for every rotation, concluding that either the sample was not perfectly aligned or the film was not well prepared. In any case, it is evident that there was LD contribution present here as well and therefore, it is important to investigate modifications of the sample preparation to assess whether it can be optimised and made reproducible. There is a significant difference also in comparison with the CD data of the films that are presented in Figure 2.20 since the sharp peaks at 210 and 220 nm are either not visible or have a very low intensity. The most probable cause for the issue is that the sample was better aligned in the case of the latest experiment, poorly prepared though, leading in slightly different signatures after each rotation. Samples must be always inspected prior to measurements to ensure that the surface is thin, smooth and even, in order to minimise the LD effect.[117] Therefore, the films from both samples were inspected using the ZEISS Stereo microscope.

In Figure 2.23(left), the film for the AB grid (Protocol F1) appears to be uneven with some crystallised areas and the different layers of samples that were applied for the sample preparation are visible, which is referred to as "coffee ring effect". This effect is not desirable since the sample is not homogeneous and it can affect the signal. Another film needed to be prepared for the AB grid (Protocol F1) sample in order to continue with the experiments. In contrast, the film for the AB grid formed with the protocol F2 in Figure 2.23(right) seems to be more homogenous and smooth, even though the film was formed by 3 layers of samples. This difference in the creation of these films may also explain the variations in the intensity of the data presented in Figure 2.20 for AB grid

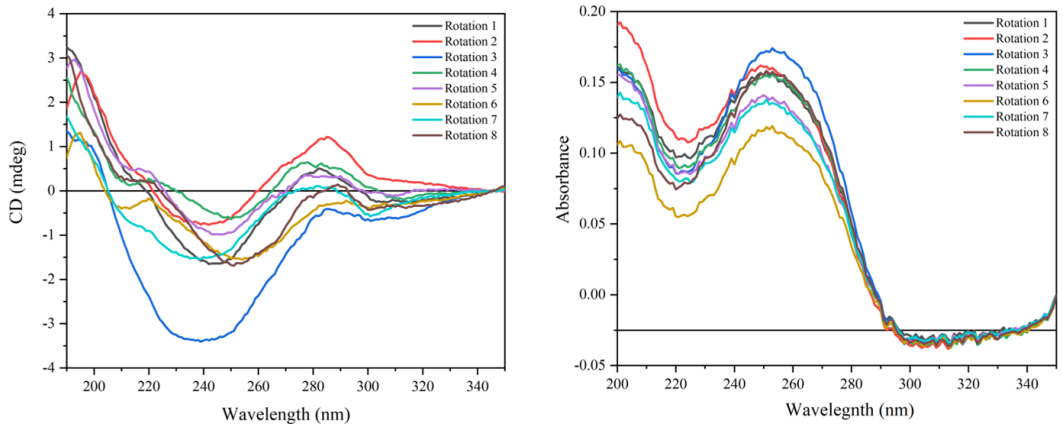

Figure 2.22: **Left**, CD spectra of AB grid film (Protocol F1), upon 8 rotations by 45° clockwise. **Right**, Absorbance spectra of AB grid film, upon 8 rotations by 45° clockwise.

protocol F1 or F2. It is essential, challenging though, to evenly deposit the same sample concentration across the same area each time in order to get reproducible results.

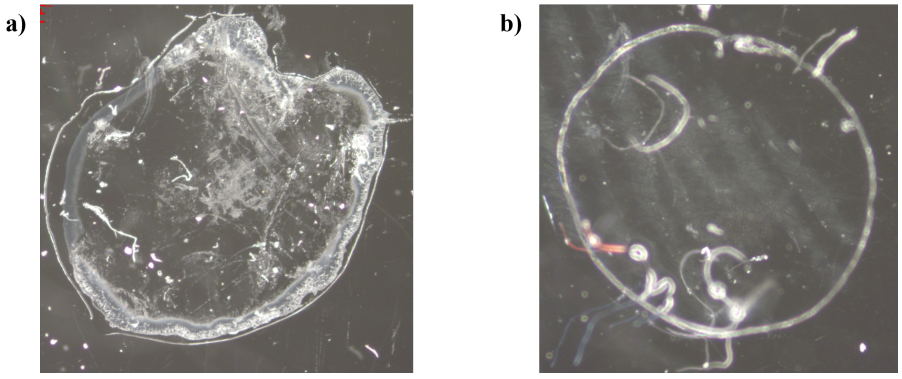

Figure 2.23: Images as captured using ZEISS Stereo Microscope. **a)** Film of the AB grid (Protocol F1). **b)** Film of the AB grid (Protocol F2).

2.2.7.3 CD *imaging*

B23 is the only beamline for SRCD that generates highly collimated and high-photon flux microbeam for chiroptical spectroscopy. This is a particular advantage since it makes possible the use of smaller amounts of samples, which is beneficial in the case of proteins. In the meantime, exploiting the bespoke vertical sample chamber equipped with motorised XY stage (Figure 2.24) enables an unprecedented, spatial resolved CD spectra at 50 μm spatial resolution can be recorded, something unattainable with bench-top CD instruments.[73] The focused microbeam in combination with the vertical chamber contribute in the investigation of oriented solid samples (films) in order to scan specific regions of the sample area and create a map of the film, which could help to spot any molecular and supramolecular changes. It can also give information about the uniformity of the film and find any uneven or defective zones. These areas can be then excluded

from the spectra something that is not applicable when larger beam sizes are used. This method was first introduced by Zinna *et al.* and it is referred to as CD imaging (CDi).[98]

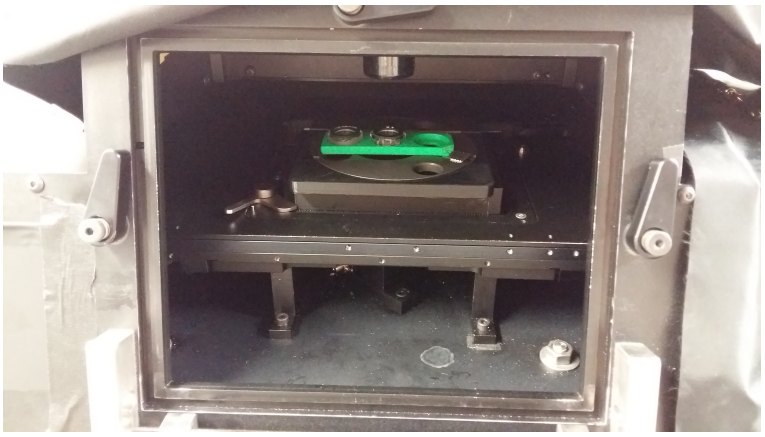

Figure 2.24: Unique B23 vertical sample chamber equipped with a motorised XY stage (Prior Stage) for imaging sample films.

A series of preliminary measurements were performed using the CDi method in order to investigate the uniformity of the films and their CD spectral features than can be used as sample signatures. Thin films of the CTB tile, AB grid (Protocol F1), AB grid (Protocol F2) and the AB grid with the protein, deposited on fused silica substrates were measured. For each film, a grid array was designed and ran in order to compose the maps of the films. The spectra of these systems were subtracted from the spectra of the corresponding fused quartz substrate windows.

CTB tile

The SRCD signature of the CTB tile film was compared with that of the solution-based sample. The absorbance was also recorded and the CDi method was used for both CD and absorbance measurements. 16 CD spectra were recorded by scanning a 4 x 4 grid array at 700 μm intervals, which are presented in Figure 2.25. The positive and negative SRCD bands were visible, but shifted to about 285 nm and 230 nm, respectively. The shape of these spectral features were slightly different than that of the B tile measured in solution (Section 2.2.7.1), since it is noticeable that the peaks at 220 and 210 nm are again sharper than those in solution.

In general, the signal for the single DNA tile is weak, as expected. All the 16 CD spectra look fairly similar, with a small difference in intensity. This can be justified by the absorbance data and the fact that the film was not homogenous in terms of concentration. It is evident that the absorbance spectra also show some variations. In order to determine the CD signature/absorbance combination, the data were processed in a multi-panel, presented in Figure 2.26.

The absorbance data are changing significantly between the scans and the most of them are below 0.1 or with no absorbance at ~260 nm. The CD signal is

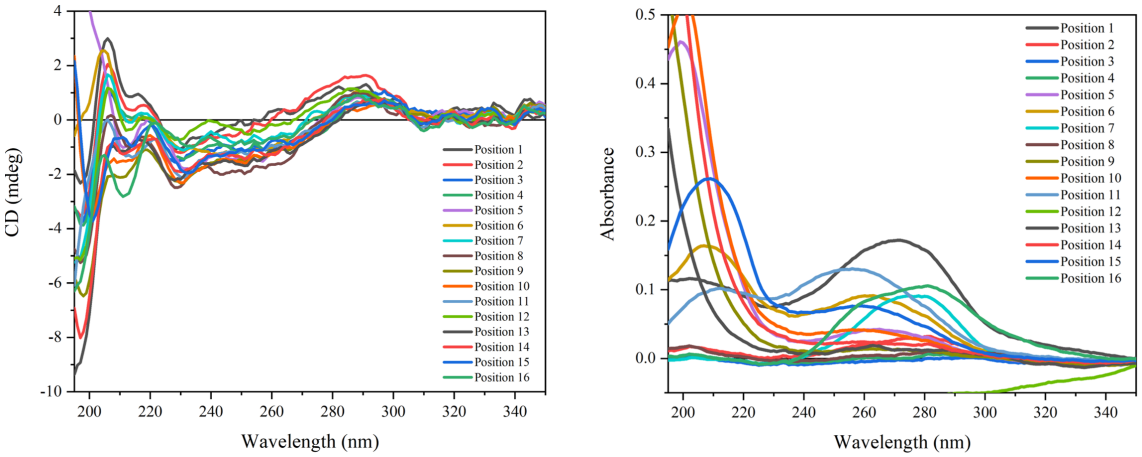

Figure 2.25: **Left**, SRCD data of the film for the CTB tile, mapping a 4 x 4 grid array of 700 μm^2 step size. **Right**, Absorbance data of the film for the B tile, mapping a 4 x 4 grid array of 700 μm^2 step size.

present even without UV absorbance which is confusing, as either presence of UV absorbance or absence of CD signal is expected. Hence, it would be useful, if both the UV absorbance and the CD intensity magnitude were increased, in order to ensure that we record above the noise and the results are corresponding to the desirable sample.

AB grid (Protocol F2)

In section 2.2.7.2, the films of the AB grid annealed with both protocols F1 and F2 were investigated using the Chirascan. It was concluded that the AB grid (Protocol F1) film was not well-prepared in comparison with the AB grid film (Protocol F2). Therefore, the AB grid film (Protocol F2) was further investigated using the Linkam tower sample compartment unit that has a camera that can be used to look at the sample guiding a more precise selection of desirable areas to be scanned.

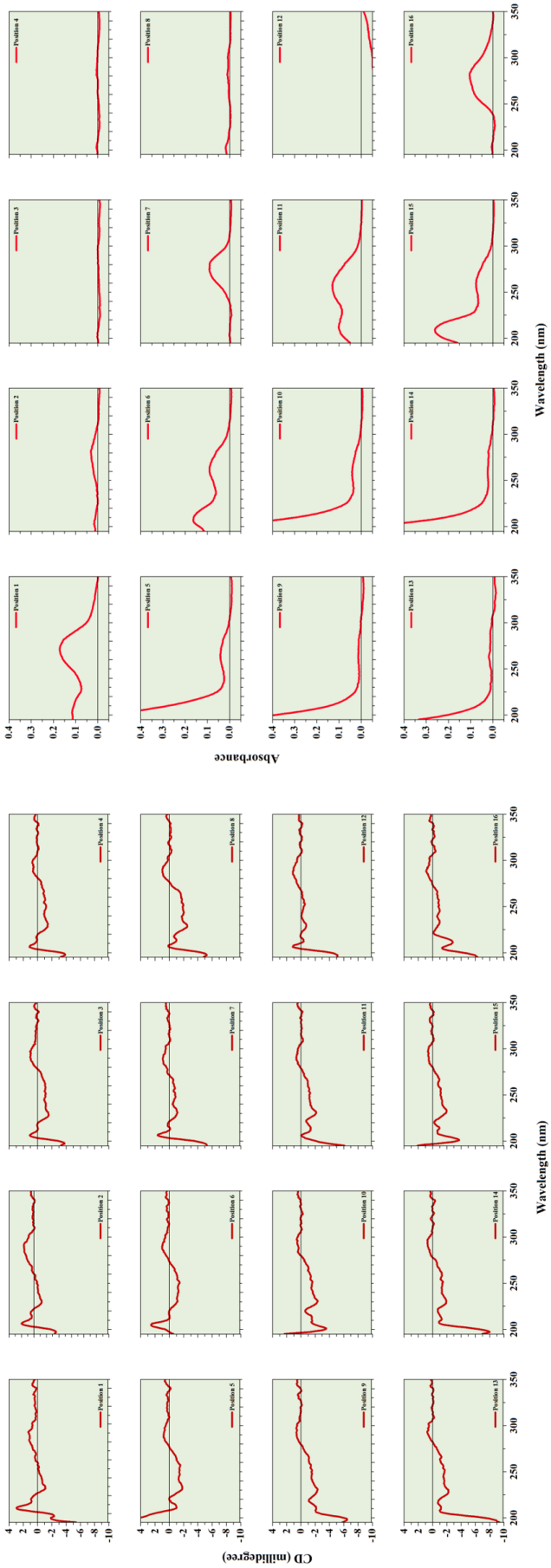

Figure 2.26: Multi-panel showing on the **left,** SRCD data of the film for the CTB tile and on the **right** absorbance data of the film for the B tile, after mapping a 4 x 4 grid array of 700 μm^2 step size.

9 CD spectra were collected for this sample as a result of mapping a 3x3 grid array at 500 μm interval steps. The same film was then rotated by 90° clockwise and the results for both cases are presented in Figure 2.27. In addition, an image of the film was captured using the camera of the chamber as shown in Figure 2.27. The film is enclosed in the dashed circle and the position of the irradiated 3x3 sampling grid is also presented in order to demonstrate the CDi concept.

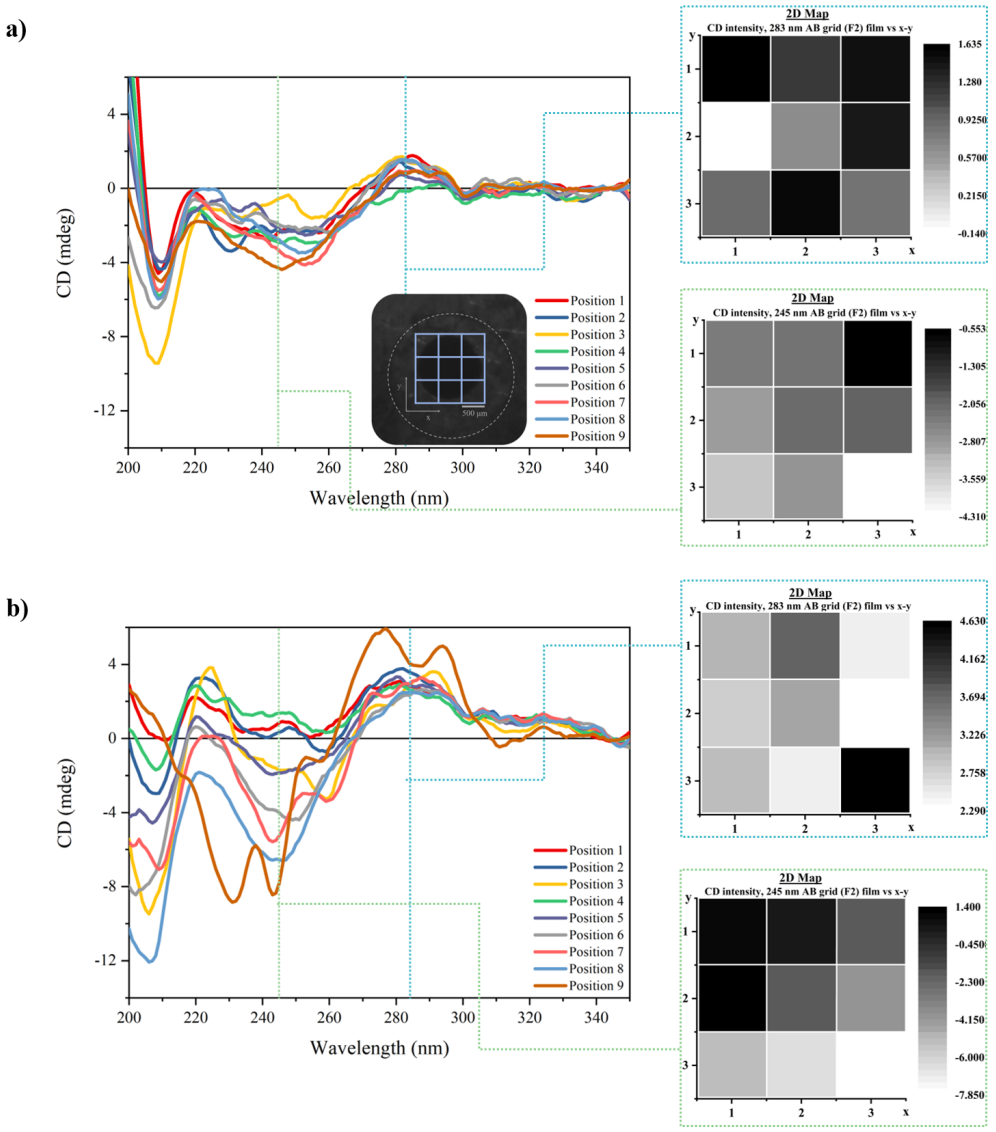

Figure 2.27: a) SRCD data of AB film (Protocol F2) after scanning 3x3 grid array at 500 μm interval steps including also an image of the film as captured with the chamber's camera, showing the irradiated 3x3 sampling area. The 2D maps were generated for the 283 and 245 nm, respectively. **b)** SRCD data and 2D maps (283 and 245 nm) upon rotation of the sample specimen by 90° rotation clockwise.

Before the film rotation, the sample looks homogeneous since all the CD signals give a similar signature for every position scanned, even though there are some small variations in the intensity of the negative peak at 245 nm. The DNA CD profile is similar to the one recorded with the bench-top instrument though weaker

(Figure 2.27). As soon as the film was rotated by 90° (without removing it from the stage holder to ensure that the grid area rotated around its centre), the variations in the intensity of the peak at 245 nm were increased along with the overall spectrum profile. The CD signal for the positive peak is the same in all areas except for the last one where the beam probably hit the edge of the film, since the whole signature does not match with the dsDNA.

Unfortunately, the absorbance was not recorded as at the time of the measurement this was not possible for the CD*i* setup. By missing those data, it is not possible to find out if those intensity variations are related to the concentration or they are due to conformational changes of the AB grid. The film formation of the same sample was attempted during a different beamtime, but unluckily it was poorly prepared and it did not produce any distinguishable signal from the noise.

AB grid (Protocol F1)

10 μL of AB grid (Protocol F1) sample was deposited on the fused silica surface and dried slowly under a stream of nitrogen in order to form a new film for CD*i* measurements. 16 CD and absorbance spectra were measured by scanning a 4x4 grid array at 700 μm step size. The 16 CD*i* spectra for both CD and absorbance are reported in Figure 2.28(a) and (b), respectively. The spectra were analysed by processing the data into distinct 2D colour maps, which were generated from the CD intensity at fixed wavelengths at 280 nm and 240 nm and from the UV absorbance intensity at 263 nm.

The shape of the spectra looks similar to the one measured with the bench-top CD instrument. Two main peaks are visible, a positive one at 280 nm and a negative one at 240 nm, which are the characteristic peaks of dsDNA in the B-form. However, the various CD*i* 2D spectra were different from each other which is better highlighted by the colour scale. Each spot has a different peak in terms of intensity but qualitatively they are quite similar. This indicated significant inhomogeneity in the film due to differences in concentration, aggregation or layer thickness. Variations in the SRCD signals can be explained by observing at the absorbance spectra. In this case, the concentration of the sample was not uniform at the surface of the film. According to the 2D colour map, the film's concentration is higher at the top-left corner of the grid array.

The CD and absorbance data were also analysed and presented in a square grid map (multipanel), which is shown in Figure 2.29, in order to inspect the CD and the UV absorbance for every spot scanned more carefully and to ensure that they are in accordance. Nevertheless, according to the grid map, it seems that in the parts with high absorbance, the CD signal is not so strong, even though the opposite would be expected. The answer to this problem could be the fact that the

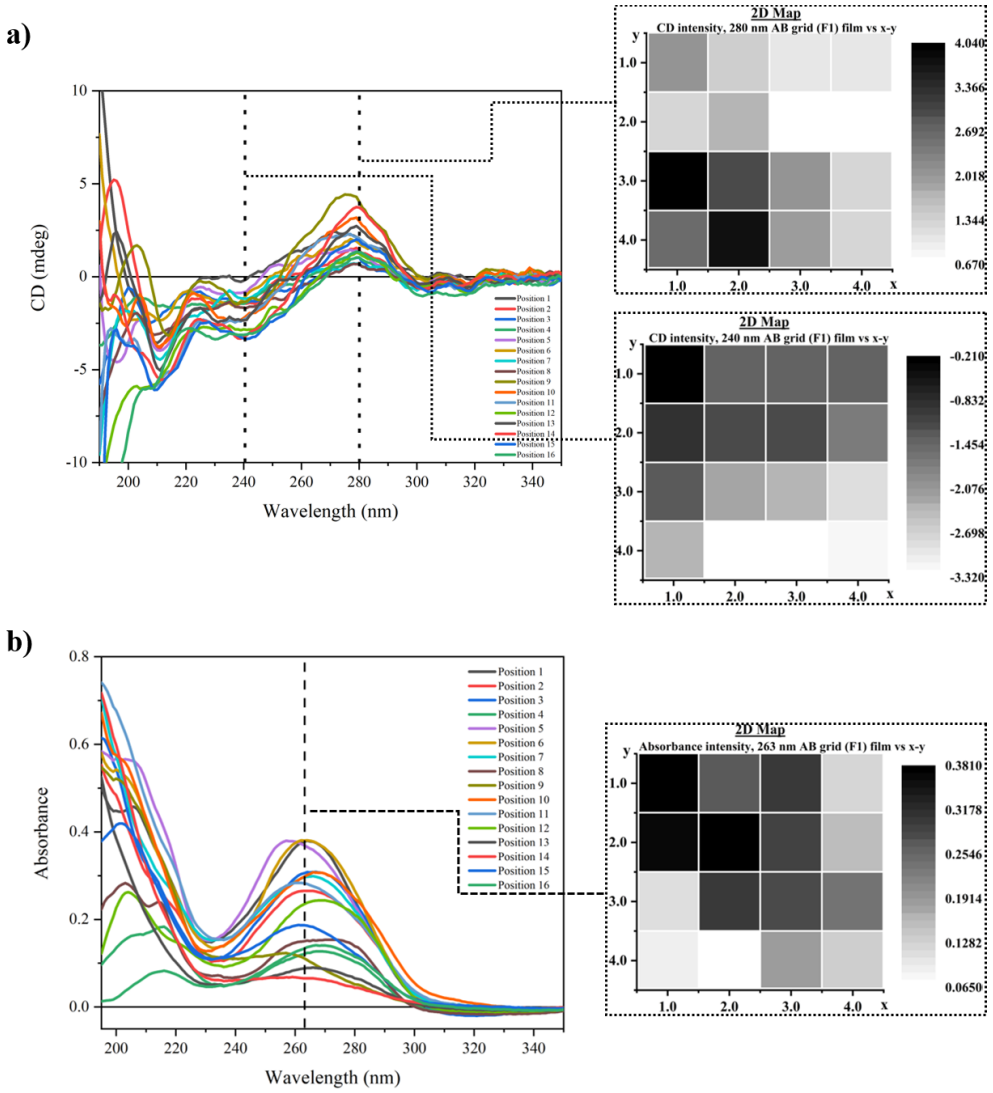

Figure 2.28: a) SRCD data of AB film (Protocol F1) after CDi experiment by scanning 4x4 grid array of 700 μm. The 2D colour maps for the fixed wavelengths 280 and 240 nm are also presented. **b)** Absorbance data after CDi experiment by scanning 4x4 grid array of 700 μm and a 2D colour map at 263 nm.

CD and the absorbance CDi measurements were not running at the same time, but they were two different runs. Even though the disk was not removed from the stage and the coordinates of the starting position were known and selected for the absorbance run, the beam might have been positioned to a shifted starting point leading to slightly different areas to be scanned between the two runs. Therefore, it was very important to be able to measure CD and absorbance simultaneously, a feature that was only supported in software later during the project. In case the disk is either removed, moved manually or even when there are issues with the automated movement of the stage, it is challenging to find the exact same position of the beam in the sample and with 50 μm accuracy, every small change counts for the imaging.

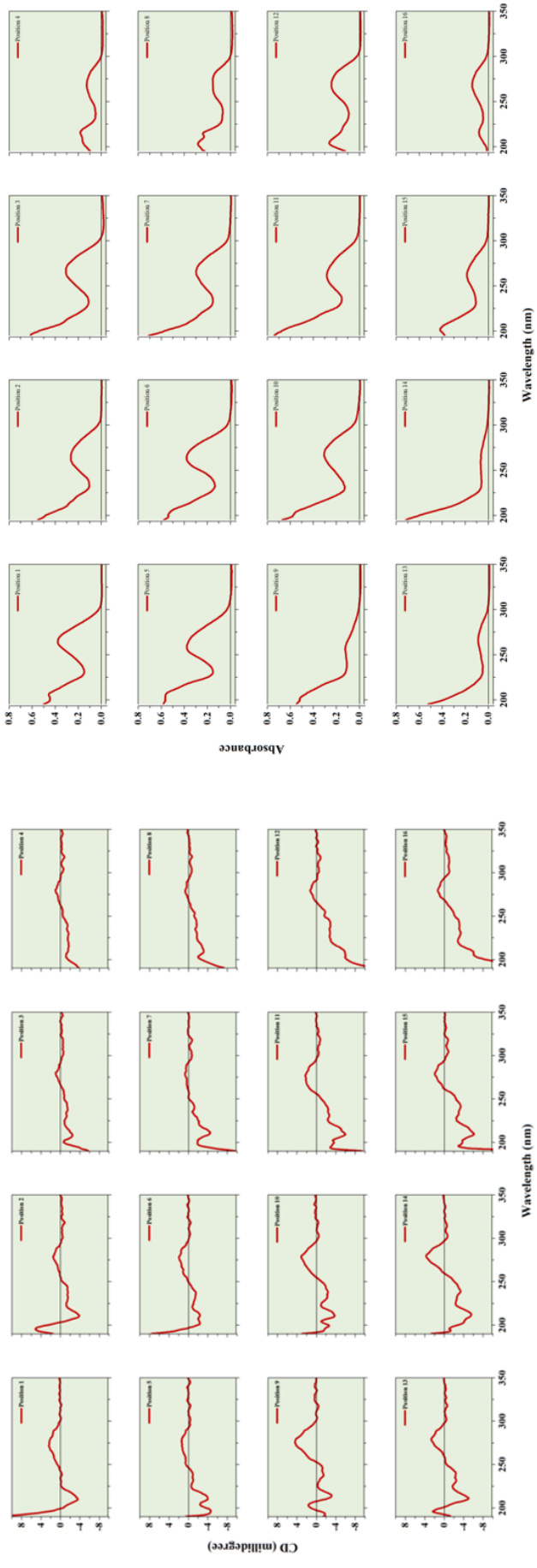

Figure 2.29: Multi-panel (square grid map) showing on the **left,** SRCD data of the film for the AB grid (Protocol F1) and on the **right** absorbance data of the film for the AB grid (Protocol F1), after mapping a 4 x 4 grid array of 700 μm^2 step size.

However, the CD and SRCD signals do not only depend on the concentration, but also on the structure of the analysed molecule. In order to determine whether the film is homogeneous in terms of structure, the measurements of SRCD and UV spectra were used to analyse the data in terms the g-factor (anisotropy factor) which is usually used for the characterisation of secondary structure of proteins in solutions and films.[119,120] G-factor reveals immediately any change in the homogeneity of the adopted supramolecular structure at the substrate surface. This is because the g-factor is the ratio between the CD in ΔA units and the UV absorption in A unit. In this way the g-factor is independent of concentration and pathlength and it is related to the thickness for films in the solid state. This is more accurate though if the CD and UV are measured simultaneously.

CD and UV-absorbance spectra are converted to dimensionless g-factor spectra by dividing the CD by the UV absorbance at each wavelength. The spectra are presented in Figure 2.30. The variation of the signals is obvious and interestingly, more negative peaks are visible at the range of 300-320 nm. The film does not appear to be uniform in terms of thickness. Therefore, the g-factor spectra of Figure 2.30 are indicative of non-homogeneous supramolecular structure, confirming the image captured with AFM which showed a non-uniform structure.

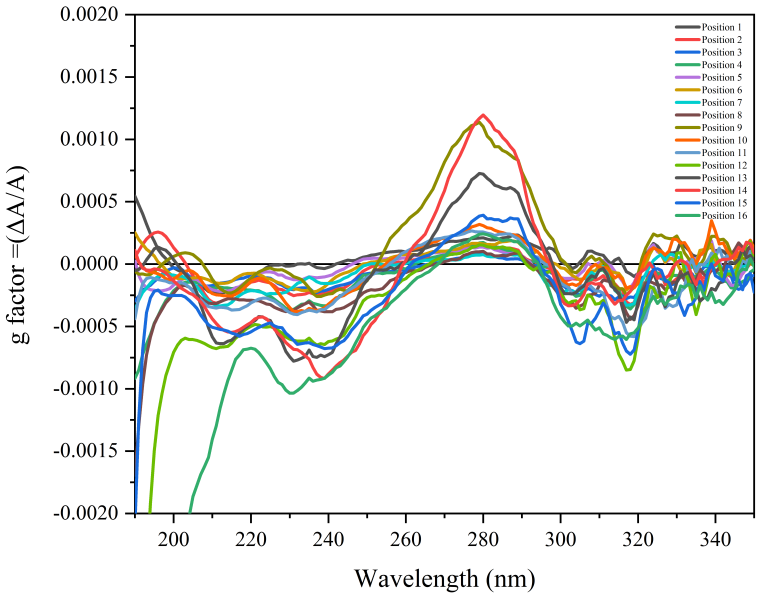

Figure 2.30: G-factor spectra of AB grid (Protocol F1) in film.

AB grid-GFP

The next film tested with the CD*i* method was the AB grid with attached GFP protein. The sample of the AB grid with the attached GFP was prepared as described in Section 6.12 and then the film was prepared by adding 10 μL of this sample on a quartz surface and dried by purging nitrogen gas. The AB grid used in this sample was formed with Protocol F1 and then hybridised with GFP. The

AB grid Protocol F2 was also investigated with GFP attached, but unfortunately none of the films produced a revealing CD signal, therefore they are not included in this thesis.

The SRCD signature and the UV absorbance of the film of the AB grid with the GFP were recorded the same way as the AB grid (protocol F1) itself. 16 CDi spectra after mapping a 4 x 4 grid array of 700 μm step size are presented in Figure 2.31.

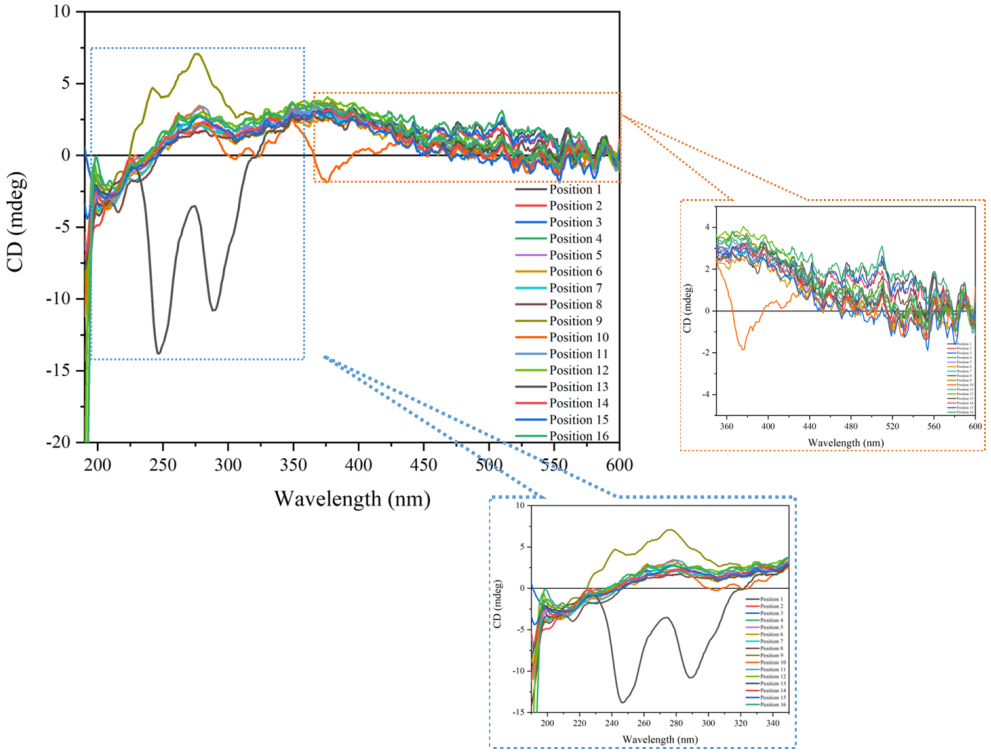

Figure 2.31: SRCD data of the film for the AB grid with GFP attached after mapping of a 4x4 grid array of 700 μm step size. The two SRCD graphs on the right are zoomed in a scale of 190-350 nm and 300-600nm, respectively.

Variations of SRCD signatures can be observed, with two scans exhibiting a completely different behaviour probably due to spectral artefacts. The positive peak at 280 nm is broader and has a lower intensity than the one in the previous SRCD spectra of the AB grid itself. On the other hand, the negative peak at 240 nm is almost invisible. The GFP absorbs at 495 nm according to the UV-Vis spectrum obtained in section 2.2.4 (Figure 2.11), hence a peak should appear in this wavelength range. The GFP is a beta-barrel and the characteristic peaks for this family of beta-sheets are a positive one at ~217 nm and a negative one at ~195 nm. However neither peaks are clearly shown in the following spectra that were consistent with spectral artefacts.

According to the UV absorbance data in Figure 2.32, the absorbance also varied, leading to the conclusion that the film was not uniform in terms of concentration.

A peak at ∼510 nm was visible and it could be a shifted peak for the absorption of the GFP (the characteristic one is at 495 nm), but it was too weak (absorbance below 0.1) in the visible region. Therefore, unless the UV bands are stronger, this measurement is not successful. The absorbance needs to be 0.3 or above in order to ensure an acceptable signal-to-noise ratio. The weak signal of the UV absorption of the GFP was consistent with the fact that there were no peaks visible in the SRCD spectra.

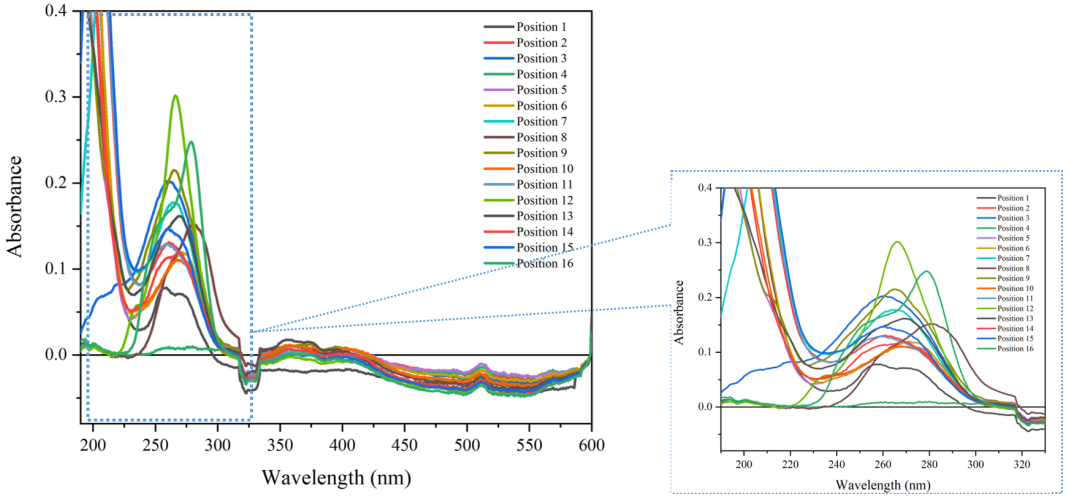

Figure 2.32: UV absorbance data of the film for the AB grid with GFP attached after mapping a 4x4 grid array of 700 μm step size. The graph on the right side is the zoomed graph data at the DNA wavelength range (characteristic peak of DNA at 260 nm).

2.3 Discussion and conclusion

In this chapter, the 2D DNA nanogrid was studied as a template for protein attachment in a precise orientation. This was possible through well-defined protein-DNA conjugation, achieved via click chemistry on the engineered proteins, such as GFP protein. One of the first challenges of this DNA nanostructure was the AFM imaging, since it was not possible to get a clear image of the 2D DNA nanogrid structure. Fortunately, some grids were visible, allowing the analysis of the system in order to ensure that the distances between the tiles were correct, which was then confirmed. In addition, when the DNA nanogrid with GFP attached was imaged, some protein dots were presented and the analysis showed that the distances between one another are in correspondence with the literature when they are attached to the tiles. However, it cannot be concluded that all the proteins observed in the images were bonded to the tile. In most cases, other features seemed to interfere with the imaging. There was an attempt of purifying the system in order to remove unwanted free tiles, free proteins or smaller structures in

general using Amicon filters, but the nanostructure being very big lead to filter blockage. Another solution would be the purification with agarose gel, which was not possible since the nanogrid sticks on top of the gel and it cannot move further due to its size.

Some preliminary SRCD experiments of the DNA nanogrids with and without protein attachment were performed both in solution and solid state as thin films using B23 beamline. In the experiments under dry conditions (films) a new method was introduced which is called CD *imaging* and it gives the opportunity, in combination with the small size beam available at B23, to scan different areas of the film resulting into a more detailed image about the desirable sample. The biggest challenge of the CD*i* method was to prepare smooth and homogeneous films, in order to avoid any artefacts that could affect the CD signal of the sample. All the samples investigated in this chapter were not very uniform, nevertheless, they produced some CD spectra consistent with dsDNA largely of B form. The films were inhomogeneous not only in terms of concentration, but also in terms of supramolecular structure, according to the g-factor analysis that were consistent with the AFM data. Despite the many attempts, the characterisation of the attached GFP protein to the DNA origami by B23 CD*i* remained elusive. The identification of the critical steps in the preparation of these protein DNA constructs revealed that fluorescence spectroscopy at high spatial resolution could be the solution for this problem as measurements with much smaller concentration used for this investigation are required. This instrument is not available yet, and B23 could be used for this purpose in the future.

Due to the challenges presented in this Chapter (and described below), the 2D DNA nanogrid was proved to be difficult for the purpose of this project. Hence, a new DNA nanostructure with a different protein than GFP will be discussed in the following chapter.

Chapter 3

DNA Double-Crossover array for protein attachment

In the previous chapter, the DNA nanostructures were introduced as templates for the attempt to immobilise a GFP protein in a precise orientation through well-defined protein-DNA conjugates, achieved via copper-free click chemistry with the engineered protein. This system was then used for SRCD experiments in order to obtain CD spectra of proteins where precise alignment is important. The first DNA nanostructure studied was the 2D DNA nanogrid which was proven to be difficult to form and showed a uniform structure resembling the expected nanogrid as reported in the literature. The GFP was chosen as a proof-of-concept protein for hybridisation with the 2D DNA nanogrid, which, failed to show detectable SRCD spectra.

In this chapter, the same recipe is followed, with the disparity that a new nanostructure is presented, as well as a new proof-of-concept protein, which have the potential to overcome some of the challenges encountered in the previous chapter. The DNA nanostructure belongs to the double-crossover (DX) molecules family containing two crossovers connecting collinear duplexes and upon successful formation, it will be used as a template for aligning an α-helical protein. In the work carried out in this chapter aims to address the challenges faced before with the following aims:

1. Synthesis and imaging of the DNA double-crossover with both liquid and dry AFM
2. Attachment of α-helical protein (modified with ssDNA) to the DNA nanostructure and investigation of the system for successful hybridisation

3. Formation of suitable thin films from the DNA nanostructure with and without attached protein for CD *imaging* experiments in order to identify conformational changes due to the hybridisation of the protein

3.1 DNA Double-Crossover (DX) Array Design Overview

The second DNA nanostructure system was the double-crossover (DX) array, which was first introduced by Seeman *et al.*.[103,121] The DX-array is fabricated from two small tiles that consist of two double helices aligned side by side, with strands crossing between helices, yoking them together. The DX tiles have an antiparallel arrangement of non-crossover strands and they belong to the double-crossover, antiparallel, even spacing (DAE) class of DX molecules which has an even number of half-helical turns between the crossover points (4 half-helical turns is 21 base pairs). In this type of tiles, three strands are found in both helices and two strands do not cross over (Figure 3.1(a)). In this project, a similar DX system of the one used by Rusling *et al.* was adopted, as shown in Figure 3.1(b).[122]

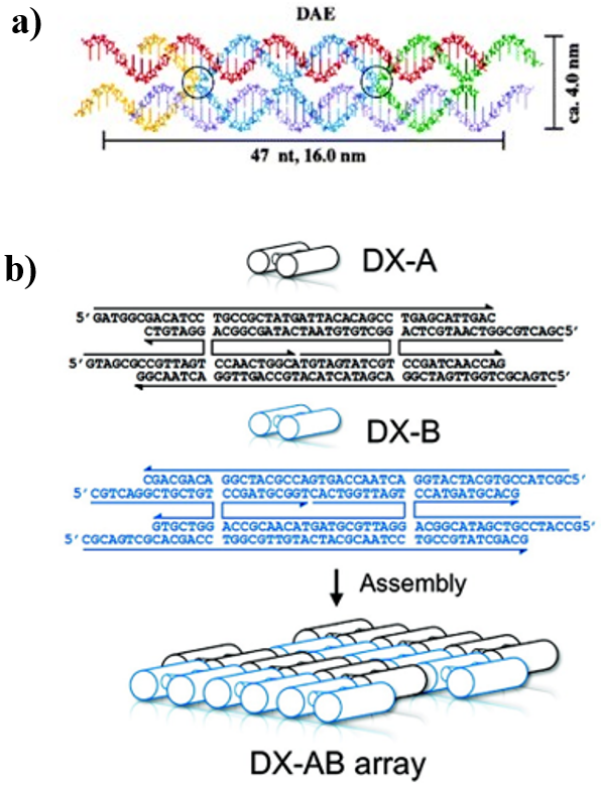

Figure 3.1: a) Model structure of DAE class of DX molecules. Adapted from[103] **b)** DX tiles and arrays. The two tiles can be mixed in equal amounts to generate the DX-AB array. Adapted from[122]

The DX tiles are programmed to contain single-stranded sticky ends with unique sequence at their four duplex termini. The association of the two DX molecules

using their complementary sticky ends generates the DX-array. A schematic representation of how the DX tiles are arranged in order to fabricate the DX-array leading to a pattern of stripes is presented in Figure 3.2.

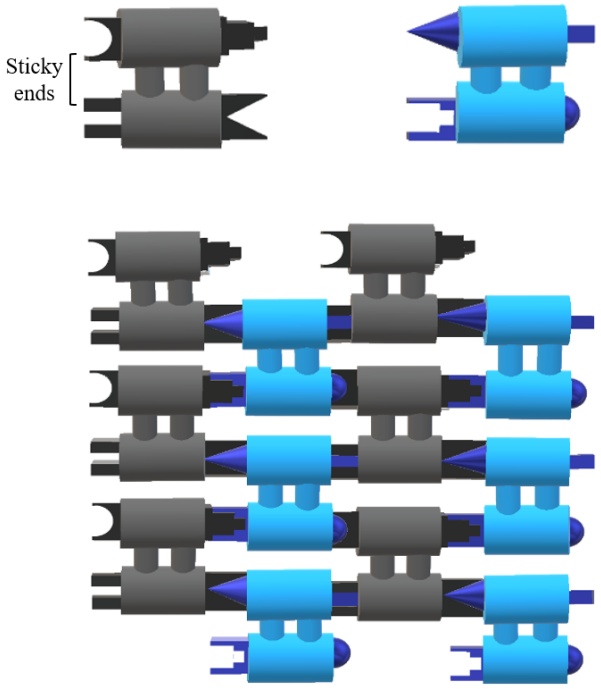

Figure 3.2: The DX tiles (grey and blue) that generate the DX-array. Their sticky ends are presented as oddly shaped adaptors which interlock to specific pieces of the other tile in order to form the bigger DX molecule

As discussed before, the DNA nanostructures are utilised as a scaffold to hold and precisely orient proteins. The cross structure designed by Hao Yan and his group had an extended ssDNA in one of the CT tiles to operate as an anchoring point for the protein attachment. However, Rusling's DX system did not include one of this. The modification of a DX tile with an extended ssDNA was therefore essential. There are many different approaches to this modification, however it is often easier to modify the central strand of one of the tiles.[123,124] Hence, the middle strand involved in the original DX-A tile was modified by extending both the 5' and 3' prime of the strand after shifting the 5' prime along by three bases, in order for the helical turn to face outwards, away from the extending lattice. The extended ssDNA was the same as the one used at CTA tile of the 2D DNA nanogrid in Chapter 3. For increased stability, the initial part of the protrusion would be a complementary double helix, thus forming a complete circular strand of DNA. This would enable the extending strand to remain as vertical as possible and would, therefore, aid the attachment of the modified protein. The initial part of the protrusion was also identical to the 2D DNA nanogrid. An additional loop of two thymidine nucleotides was introduced to avoid increased stiffness in the tile. In the end, the new demonstrated tile would include the crossovers, as well as one

double-helical region (junction), like a bump at the top of the tile, which ends up as ssDNA and serves as the hybridisation point of the protein.

However, in the case of the CTA tile of the 2D DNA nanogrid, the extended single stranded area (protein hybridisation site) was located at the 5' prime end of the strand. When the protein was modified with the ssDNA, complementary to the hybridisation site, it was introduced on the 5' prime of the DNA, and, therefore, resulted to be close to the surface of the grid (Figure 2.12) after its hybridisation to the DNA nanostructure. Using the same part of protrusion and extended ssDNA for the modified DX-A tile, the protein should be positioned closer to the surface of the array. In order to investigate whether the location of the protein on the DX-array affects the supramolecular structure, another version of the DX-A tile was prepared, which had the extended ssDNA on the 3' prime of the middle strand and therefore, the protein was positioned at a bigger distance from the tile surface. For this reason, for the rest of the thesis the DX-A tile with the ssDNA on the 5' prime will be referred to as DX-$A_{5'}$, whereas the tile with the ssDNA on the 3' prime will be called DX-$A_{3'}$, and the DX-arrays, DX-array$_{5'}$ and DX-array$_{3'}$, respectively. In Figure 3.3, the original DX-A tile along with the modified ones are illustrated.

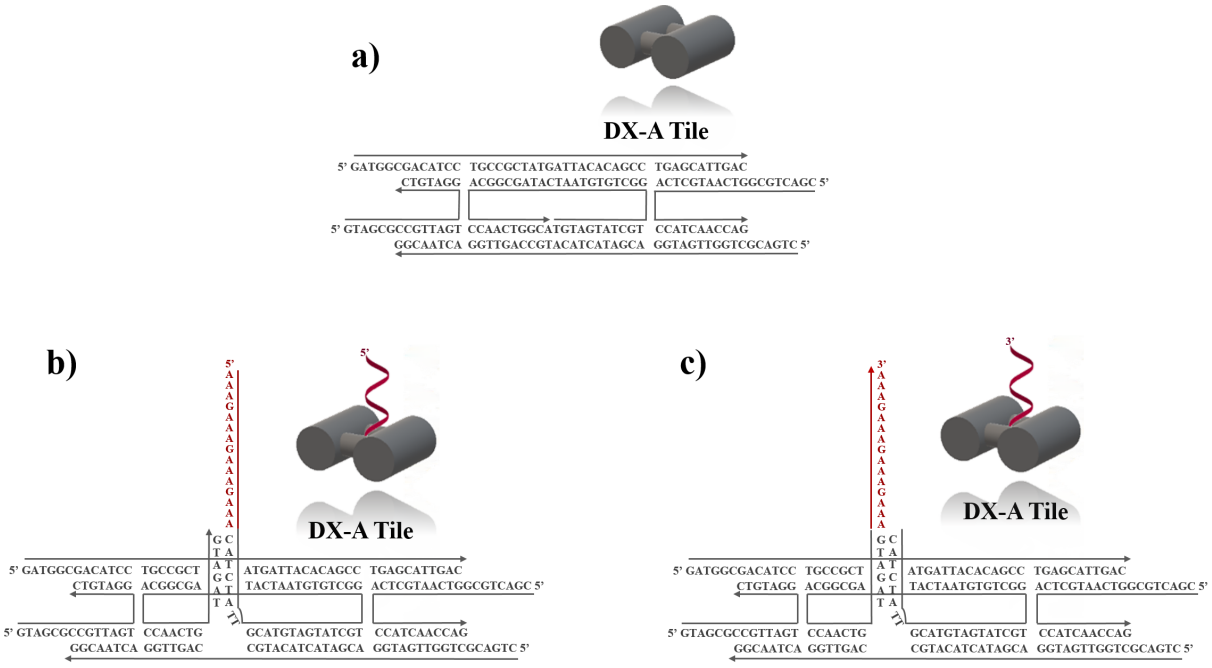

Figure 3.3: **a)** Original DX-A tile with sequences included. Adapted from [122] **b)** and **c)** Modified DX-A tiles, DX-$A_{5'}$ and DX-$A_{3'}$ with the extended ssDNA (red) either on the 5' prime or the 3' prime end, respectively.

The final schematic representation of the DX system that is used in this chapter is presented in Figure 3.4.

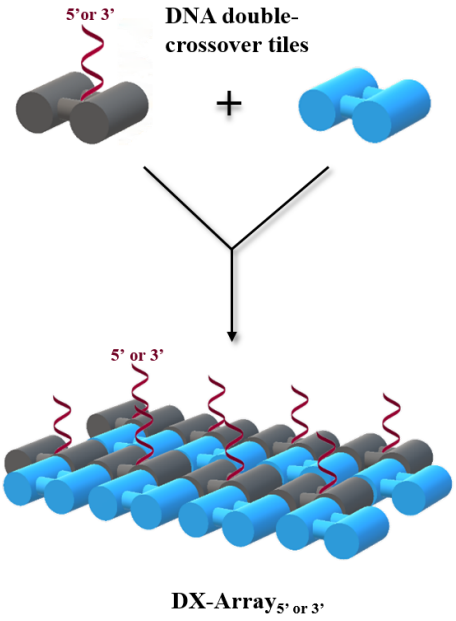

Figure 3.4: DX tiles and the DX-array. The DX-A tile (grey) and DX-B tile (blue) contain complementary sticky ends and then can be mixed in equal amounts to generate the DX-array. The DX-A consist of the extended strand (red), the anchoring point for the protein attachment.

3.2 Results and discussion

3.2.1 Synthesis and analysis of the oligonucleotides

Sets of 10 DNA strands, 5 for DX-A and 5 for DX-B tiles, were synthesised. The sequence of each oligonucleotides included in DX-A and DX-B can be found in Table 3.1. The DNA strands DX-A$5_{5'}$ and DX-A$5_{3'}$ are the modified ones, as described earlier (Section 3.1) in order to introduce the extended ssDNA (either at the 5' or the 3' prime, respectively) for the protein attachment. In Figure 3.3(b) and (c), the DX-A tiles with their sequences were introduced as schematic representation. In Figure 3.5, the DX-B tile with the sequences for each DNA strand is also presented.

Figure 3.5: DX-B tile with sequences included.

Strand	No. of bases	Sequence (5' to 3')	Strand	No. of bases	Sequence (5' to 3')
DX-A1	45	GAT GGC GAC ATC CTG CCG CTA TGA TTA CAC AGC CTG AGC ATT GAC	DX-B1	21	CGT CAG GCT GCT GTG GTC GTG
DX-A2	47	CTG ACG CTG GTT GAT CGG ACG ATA CTA CAT GCC AGT TGG ACT AAC GG	DX-B2	31	GCC ATC CGT CGA TAC GGC ACC ATG ATG CAC G
DX-A3	21	GTA GCG CCG TTA GTG GAT GTC	DX-B3	42	GGT CAC TGG TTA GTG GAT TGC GTA GTA CAA CGC CAC CGA TGC
DX-A4	31	CGA CTG CGG TCA ATG CTC ACC GAT CAA CCA G	DX-B4	47	CGC TAC CGT GCA TCA TGG ACT AAC CAG TGA CCG CAT CGG ACA GCA GC
DX-A5_5'	71	AAA GAA AGA AAG AAA CAT CTA TTG CAT GTA GTA TCG TGG CTG TGT AAT CAT AGC GGC ACC AAC TGT AGA TG	DX-B5	49	CGC AGT CGC ACG ACC TGG CGT TGT ACT ACG CAA TCC TGC CGT ATC GAC G
DX-A5_3'	71	CAT CTA TTG CAT GTA GTA TCG TGG CTG TGT AAT CAT AGC GGC ACC AAC TGT AGA TGA AAG AAA GAA AGA AA			

Table 3.1: The first 3 columns from the left-hand side present the sequences of the oligonucleotides for the DX-A tile, whereas the 3 columns from the right-hand side present the sequences of the oligonucleotides for the DX-B tile.

The DNA synthesis was carried out using the Expedite Synthesiser from Applied Biosystems and then purified following the same procedure described in Section 2.2.1. The Beer-Lambert Law (equation 2.1) was used to calculate the concentration of those oligos upon absorbance measurement using UV-Vis spectroscopy. The absorbance spectra recorded are presented in Figure 3.6. The DX-A5$_3$, is not included in the spectra since it was introduced at a later stage on in the project and it was purchased from Sigma-Aldrich.

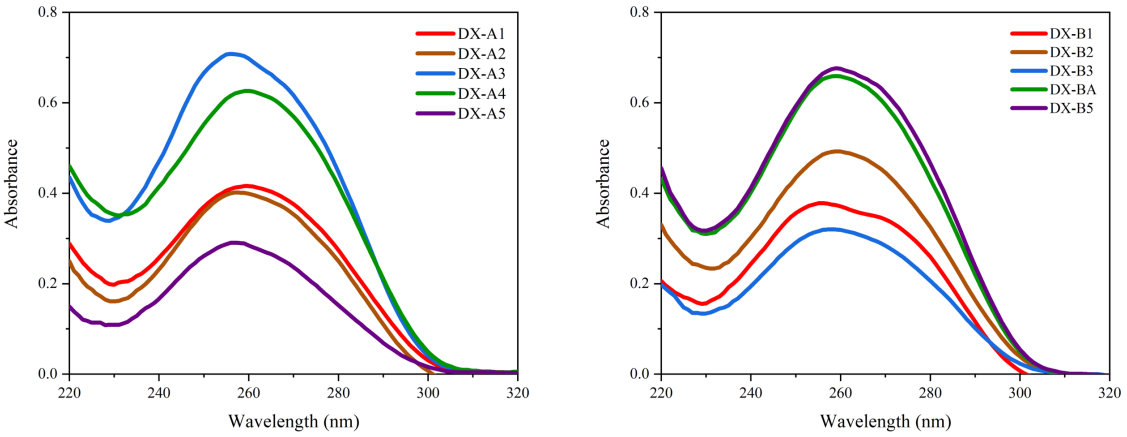

Figure 3.6: UV-Vis spectra of unmodified DNA with a peak at 260 nm for DNA. UV-Vis spectra of the strands involved in DX-A tile **(left)** and in DX-B tile **(right)**.

The number of moles of DNA was also obtained and the unmodified oligonucleotides were further analysed using mass spectrometry. The mass of each strand is shown in Table 3.2 along with the number of moles of DNA and the synthesis yield. The traces of the MS analysis can be found in Appendix B.

DX-A Tile	Calculated Mass (g/mol)	Mass Found (g/mol)	Number of moles (nmol)	Yield (%)
DX-A1	13805.9	13803.9	176.4	17
DX-A2	14519.3	14517.4	160.2	16
DX-A3	6493.2	6491.4	673.2	64
DX-A4	9450.1	9448.7	388.2	39
DX-A5$_5$'	22030.2	22025.2	71.9	7.2
DX-B Tile	**Calculated Mass (g/mol)**	**Mass Found (g/mol)**	**Number of moles (nmol)**	**Yield (%)**
DX-B1	6485.1	6483.5	360.0	36
DX-B2	9466.1	9464.6	304.2	30
DX-B3	12970.3	12968.1	144.0	14
DX-B4	14403.2	14410.3	270.0	27
DX-B5	14994.6	14992.5	270.0	27

Table 3.2: Analytical data of Mass Spectrometry for DX-A and DX-B tiles.

3.2.2 Formation of the DX tiles and array

Before the formation of the DX-array, the oligos were further purified using denaturing urea DNA gel electrophoresis (PAGE purification, Section 6.5) in order to ensure that the oligos are as pure as possible and there are no unwanted smaller sequences that might cause the formation of undesired structures. In addition, the best purification methods for oligos with a length, which overcomes 60 bases, like the DX-A5 strand (71 bases), is PAGE purification compared to High Performance Liquid Chromatography (HPLC), which is suitable for shorter oligos.

After this extra purification step, the oligos were used for the annealing of the DX-$A_{5'}$ or DX-$A_{3'}$ and the DX-B tiles and finally the formation of the DX-Array following Protocol F2 as described in Section 6.10. In order to clarify whether the individual DX-A and DX-B tiles were formed, as well as the DX-array, 1.0% agarose gel electrophoresis was carried out. Two different controls were used so that they could be compared with the tiles and the DX-array. Controls a1 and b1 consist of the strands DX-A1, DX-$A5_{5'}$ and DX-B3, DX-B4 respectively, mixed together, but not heated. Controls A1 and B1 consist of all the strands present in each tile individually, mixed together, but not heated. According to the gel presented in Figure 3.7, tiles DX-A and DX-B were obtained. The DX-array, since it is a large structure, cannot run compared to the tiles and, therefore, remained at the top of the gel. The strands consisting of Controls a1 and b1 also formed the individual tiles spontaneously, regardless of not being heated. Hence, they are visible in the same level with the band of the DX-A and DX-B tile, respectively.

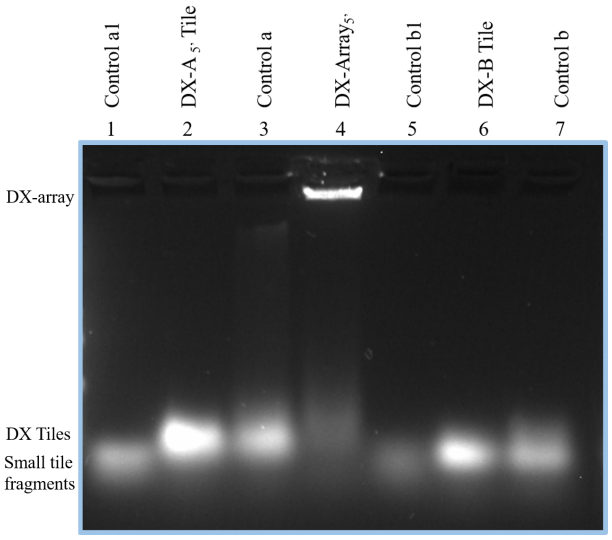

Figure 3.7: 1.0 % agarose gel. Lane 1: Control a1, Lane 2: DX-$A_{5'}$, Lane 3: Control a, Lane 4: DX-$Array_{5'}$, Lane 5: Control b1, Lane 6: DX-B, Lane 7: Control b.

3.2.3 Protein Modification- Click Chemistry

Similarly to Section 2.2.3, the protein hybridisation to the DX-array requires the protein to be modified with an ss-DNA which has the complementary sequence of the DX-A tile's hybridisation site. Therefore, the concept was that the protein would be genetically engineered with an azide which would then react with the 5'-Click-easy™ BCN CEP II (alkyne) modified DNA through SPAAC click chemistry. As discussed in the conclusion (Section 2.3) of the Chapter 2, the modified ssDNA was synthesised 18 bases long instead of 15. In this chapter, the sequence was corrected (Table 3.3) and the BCN-modified DNA was purchased from ATDBio.

Strand	No. of Bases	Sequence (5 '- 3')
5' BCN-DNA 1	15	YTTTCTTTCTTTCTTT

Table 3.3: Sequence of the modified oligonucleotide, where Y is the 5'-Click-easy™ BCN CEP II .

In this Chapter, a new protein was introduced for protein attachment to the DX-array, since the sfGFP^{azF204} did not yield the desirable results in the CD experiments. This new protein, cytochrome b_{562} (cytb_{562} hereafter), is a soluble, 12kDa α-helical protein which is expressed in the periplasm of *Escherichia coli*.[125,126] The protein is comprised of four antiparallel α-helices that wrap into a left-handed bundle and binds non-covalently a single iron heme through the axial ligands methionine (Met7) and histidine (His102).[127–129] According to the literature, cytb_{562} tolerates azF inclusion at different positions in its structure.[110] The protein samples of cytb_{562} were provided by Dafydd Jones' PhD student, Benjamin Bowen, from Cardiff University, who carried out the mutation of the protein with the azF and the click chemistry of the genetically modified protein with BCN-DNA.

Initially, two variants of the protein were provided in their homodimer form, linked together by a disulphide bond between their cysteine's SH group. The first variant was the "Short Axis" (SA) cytb_{562} with an azF mutation at position 5 and a cysteine opposite at position 104. The second one was the "Long Axis" (LA) cytb_{562} with azF mutation at position 50 and a cysteine opposite at 21. Cysteine mutation was indeded for attachment to gold surfaces as part of Benjamin Bowen's work but does not play a role here. The sequences of the mutated variants can be found at the Appendix C. The protein variants along with the positions of the mutations are presented in Figure 3.8. Both variants were mixed at a 1:2 ratio (DNA/protein) with BCN-DNA. The results of the SDS-PAGE and the UV-Vis obtained by Benjamin Bowen are shown in Figures 3.9 and 3.10.

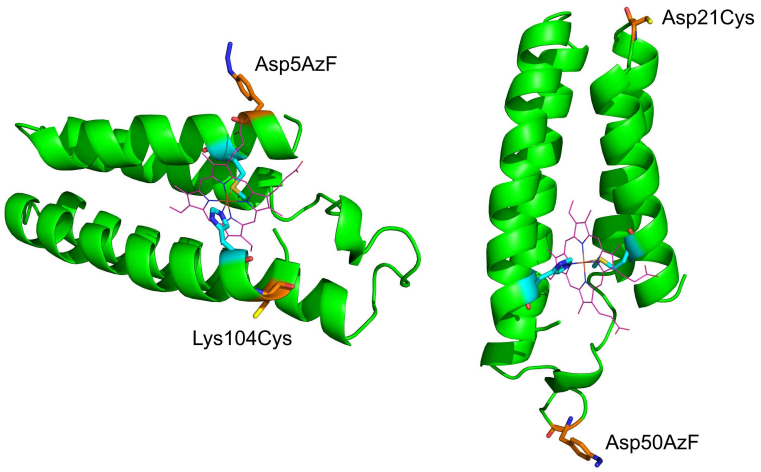

Figure 3.8: Left, short axis (SA) cytochrome b_{562} with azF mutation at position 5 and cysteine mutation at position 104. **Right,** long axis (LA) cytochrome b_{562} with azF position at 50 and cysteine mutation at position 21.

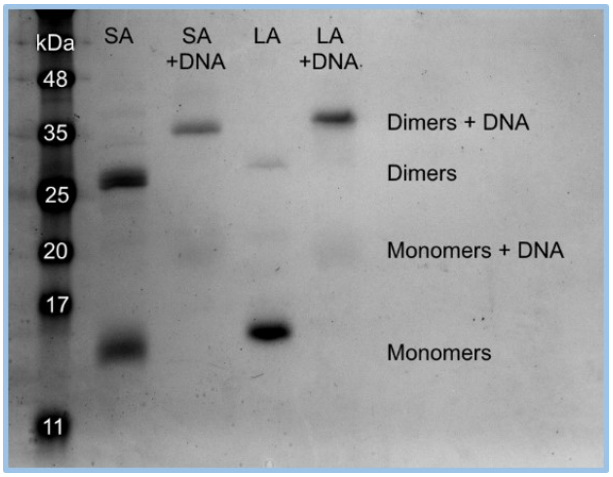

Figure 3.9: Gel mobility shift analysis of protein-DNA conjugation. The left hand lane is the molecular weight standards. Lanes represented the unmodified variants of cytb_{562} and the protein-DNA are annotated on the figure. The monomeric and the dimeric form of the variants are visible. This SDS-PAGE image was obtained by our collaborator Benjamin Bowen.

For both protein variants in Figure 3.10, the protein in the protein+ssDNA sample is a dimer and that explains its similarity in the spectrum to the SA or LA cytb_{562} dimer on its own. The difference in the intensity of heme's λ_{max} for SA cytb_{562} and SA cytb_{562}+ssDNA may be due to different concentration of samples. In the case of the LA variant, the two spectra (black and blue) overlap if one of the spectra will be shifted up or down, suggesting that those two samples have similar concentration. However, it is evident that the monomer for both variants gives slightly shifted λ_{max} at around 426 nm instead of 418 nm that the dimer has. This variation in the λ_{max} resembles the absorption characteristics of the ferric form (oxidised, Fe^{3+}) of cytb_{562} with Soret band at 419 nm and the ferrous form (reduced, Fe^{3+}) with λ_{max} at 427nm.[130] Furthermore, based on the information

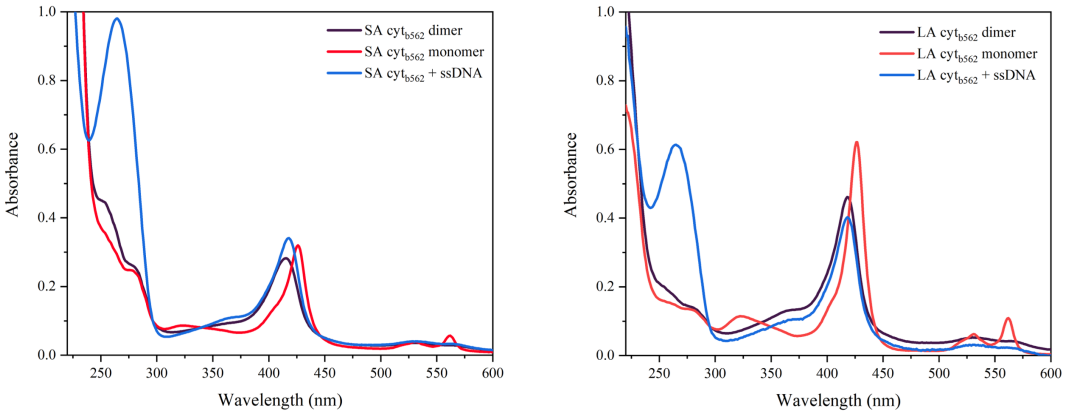

Figure 3.10: Left, UV-Vis spectra of SA cytb_{562} dimer, SA cytb_{562} monomer and SA cytb_{562}+ssDNA with a peak at 265 nm for the DNA and at 418 nm for the cytb_{562}. **Right**, UV-Vis spectra of LA cytb_{562} dimer, LA cytb_{562} monomer and LA cytb_{562}+ssDNA with a peak at 265 nm for the DNA and at 418 nm for the cytb_{562}. The UV-Vis data were obtained by our collaborator Benjamin Bowen.

provided by our collaborators, in order to go from dimer to monomer , they used a reducing agent and that explains why the monomer ended up with a red shift in the Soret band indicating that the monomer for both variants is predominantly in its ferrous form.

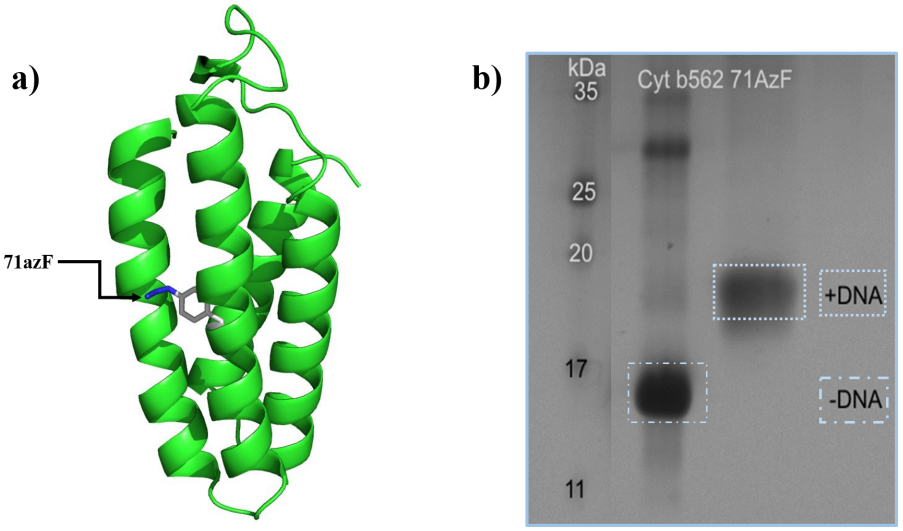

Figure 3.11: a) Short axis cytochrome b_{562} (Q71) including the location of the azF mutation on the sort axis of the protein at position 71. **b)**, Gel mobility shift analysis of protein-DNA conjugation. The left hand lane is the molecular weight standards. Lanes from left to right show bands of the unmodified and the modified variant of Q71 cytb_{562}. The SDS-PAGE image was obtained by our collaborator Benjamin Bowen.

At a later stage, one more variant of short axis cytb_{562} was added, the Q71 cytbb_{562}. This variant has the azF mutation at position 71 and it was provided as a monomer. The approximate structure of the protein variant, and the SDS-page gel taken by Benjamin Bowen are presented in Figure 3.11(a) and(b), respectively. It is not identical compared to the SA cytb_{562} (azF in position 5), but they have similar

orientation. In the gel (Figure 3.11(b)), the bands for the Q71 cytb_{562} before and after the BCN-DNA modification are visible in the expected levels according to their molecular weight.

All the variants demonstrated above were used for hybridisation with the DX-array, followed by CD experiments.

3.2.4 Protein-ssDNA attachement to DX tile and array

The proteins were mixed with the DX-array in 1:1 or 2:1 molar ratio. The DX-array with the different variants of cytb_{562} were co-assembled at 25°C overnight on thermal cycler. As mentioned earlier in this chapter, depending on which end (5' or 3') the middle strand of the DX-A tile has the recognition sequence for hybridisation with the complementary strand of the modified protein, the protein will be positioned either closer to the surface of the DX-array or in a bigger distance from the tile surface. The schematic representation of the DX-array system before and after the protein attachment is presented in Figure 3.12.

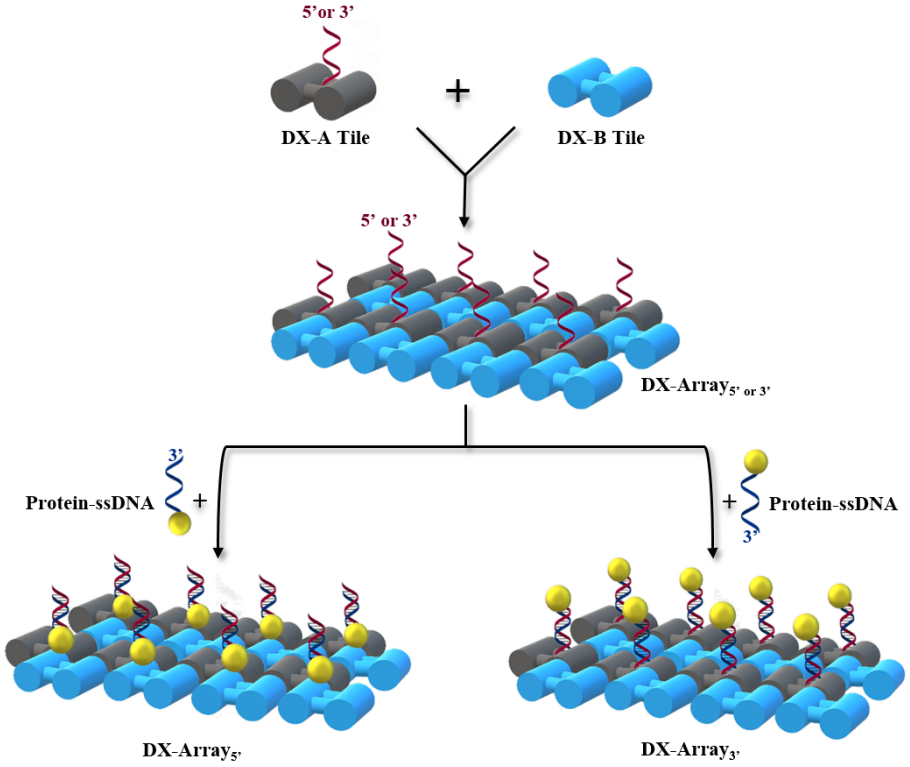

Figure 3.12: Schematic representation of the final DX-array system. DX-A tile and DX-B tile are mixed in equal amounts to generate the DX-array, which includes the extended strand on each DX-A tile (either on the 5' or 3' end of the middle DNA strand) pointing out of the plane. Finally, the protein of interest is immobilised on the DX-array either close to the array surface or on top of it.

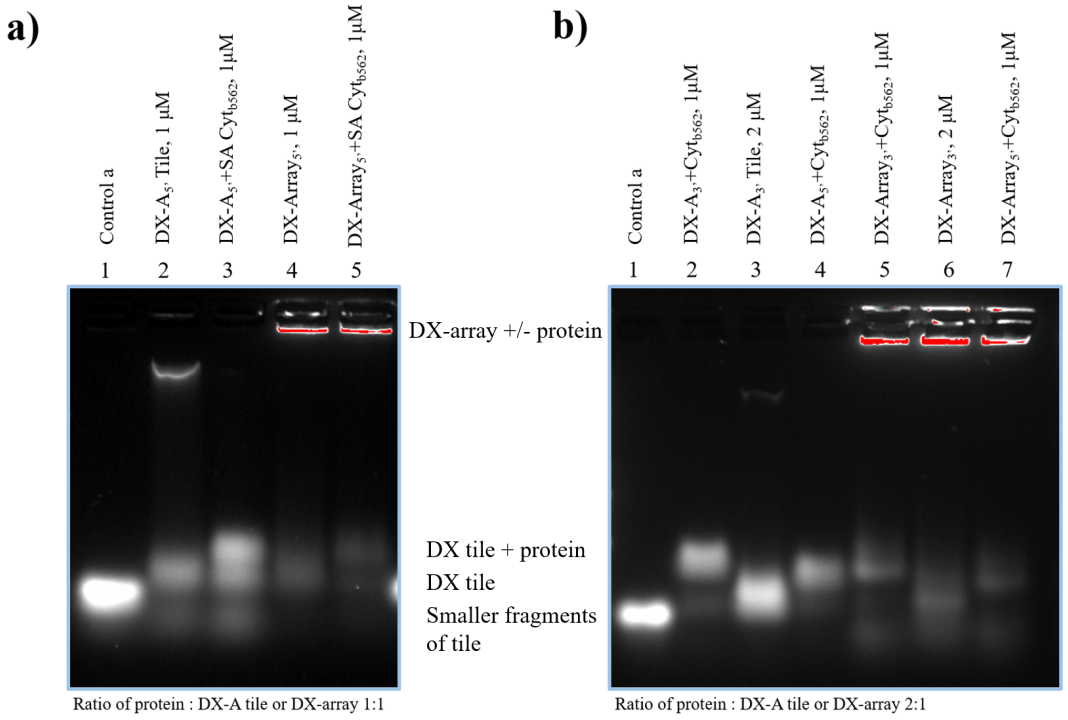

Figure 3.13: a) 1.0% agarose gel of the DX-$A_{5'}$ tile and DX-$Array_{5'}$ with and without protein (SA cytb_{562}, azF in position 5). Ration of protein:DNA=1:1. **b)** 1.0% agarose gel of the DX-$A_{3',\text{ or } 5'}$ tile and DX-$Array_{3',\text{ or } 5'}$ with and without protein (Q71 cytb_{562}). Ration of protein:DNA=2:1.

In order to confirm that the protein was attached to the tile, the protein-ssDNA was mixed with the DX-A tile in a ratio 1:1 or 2:1 in order to investigate which ratio gives less free DX-A tile and subsequently, they were visualised using 1.0% agarose gel. Control a (mix of DX-A1 and DX-$A5_{5',\text{ or } 3'}$ DNA strands) and the individual DX-$A_{5',\text{ or } 3'}$ tile were used as controls. According to the gel presented in Figure 3.13(a), which has protein-DNA ratio of 1:1, the protein binds to the DX-A tile since there is a band in a higher level for the DX-A+SAcytb_{562} compared to the DX-A. However, there is still a high portion of DX-A tile band left in the DNA-protein lane that runs as an individual tile. Some strands are also visible at the lower part of the gel, which might have fallen apart while the gel was running. On the other hand, in the gel where the protein-DNA molar ratio is 2:1 (Figure 3.13(b)), the protein seems to bind again to the DX-A tile and the amount of the free tile is less than the one with 1:1 ratio. It is also noticeable that in the case of gel B, when the protein is attached at the DX-$A_{3'}$ (far from the tile surface), it yields a band in slightly higher level than the one of the DX-$A_{5'}$+cytb_{562} (closer to the tile surface), which could be further investigated using CD spectroscopy. This will enable the comparison of the CD signals of those two different positions of the protein-ssDNA non covalently bound to the array. Finally, in the case of the DX-array with the attached protein in both gels, it is hard to determine whether the hybridisation is successful, since the DX-array hardly moves on the gel due to

its larger size. It is assumed that it is successful considering that it binds to the individual tile, but this will be further characterised and confirmed by AFM.

3.2.5 Atomic Force Microscopy of DX-Array with protein attachment

Further investigation on whether the DX-array was successfully formed, along with the accomplishment of the protein hybridisation, was performed using AFM. However, in contrast to the 2D DNA nanogrid, the DX-array was only imaged after the hybridisation of the protein on its surface, since it would be easier to spot the structure on the mica. Protein hybridisation takes place at every other tile within the DX-array and according to the literature, the repeat spacing of the proteins corresponds to 32 nm that is 94 nucleotides apart.[122] All AFM images were obtained using tapping mode both in air and in liquid of the samples on freshly deposited on cleaved mica.

3.2.5.1 Liquid AFM of DX-Array with GFP attachment

The sample of the DX-array with attached GFP (1:0.7 ratio, where 1 is 1 μM) was imaged with tapping mode at liquid AFM (FastScan Bio™ from Bruker) at Strathclyde University in Glasgow. 5 μL of sample was spotted on the mica and then, 20 μL of the annealing buffer was added on top. In Figure 3.14, it is shown that the tiles were joined together by their sticky ends creating a regular pattern of stripes, which corresponds to the data reported in the literature (Figure 3.14, b).[103]

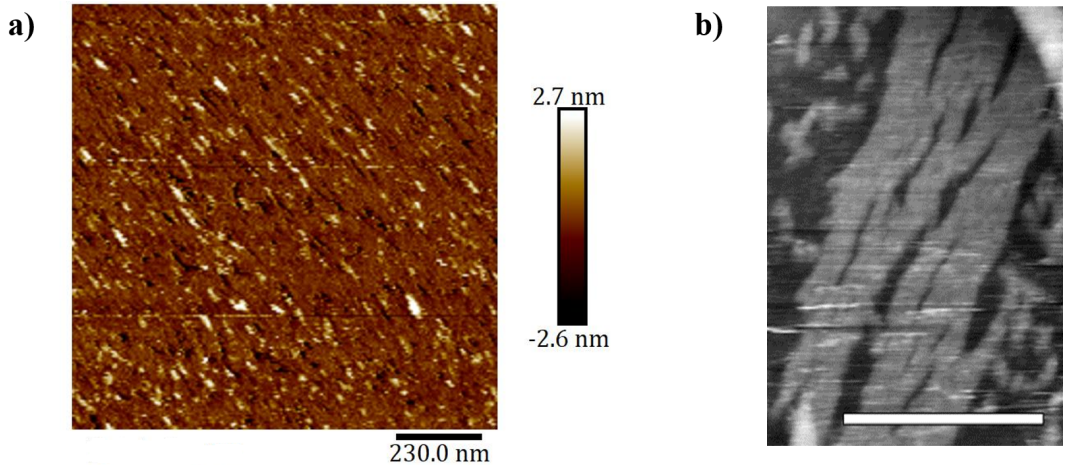

Figure 3.14: a) AFM image of DX-array with GFP attached (1:0.7 ratio) as captured wih liquid AFM, 1.2 μm scan size.**b)** DAE-O AB lattice with 33±3 nm periodicity, 500nm scan size, captured by Seeman *et al.*[103]

It appears that the attachment of the protein was successful, as bright spots aligned to the linear pattern were visible. However, the sample is too concentrated and as a result, it was difficult to perform the section analysis. Consequently, the spacing between the protein molecules and the height of the sheets could not be obtained and therefore, comparison with the literature was not possible. At a glance, though, the DX-array formation and the protein attachment to the DNA surface appear to have been performed successfully.

3.2.5.2 AFM under dry conditions of DX-Array with Cytb_{562} attachment

The sample of DX-array with attached cytb_{562} (short axis) was prepared for imaging using AFM under dry conditions. The ratio between the DNA and the protein was 1:1, with each part being 1 μM. 5 μL of sample was spotted on freshly cleaved mica surface, left for hybridisation for 3 minutes, washed with deionised water and dried under nitrogen. The sample was then imaged using tapping mode.

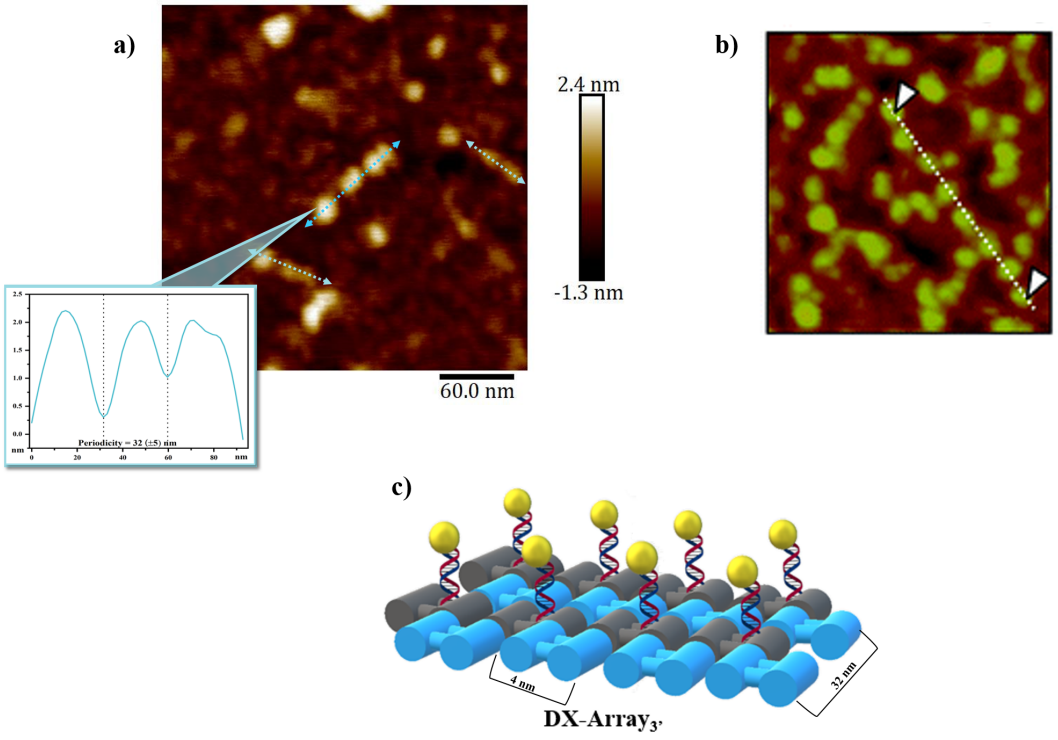

Figure 3.15: a) AFM image of 1 μM DX-array with cytb_{562} attached and the section analysis of the light blue dotted arrow showing the distance between proteins, scan area 302.6 nm **b)** AFM image of sample containing the DX-array, the bioti-TFO and an excess of SA as captured by Rusling, 300 nm scan size. Adapted from[122] **c)** The expected distances between tile to tile.

According to the results presented in Figure 3.15, clear patches are revealed on the mica surface that can be ascribed to DNA lattices with attached cytb_{562}. The section analysis of these patches showed a height of 2.0 ($\pm$0.2) nm, which

is in good agreement with a monolayer of DNA and it is consistent with the literature.[103,122] The average spacing is 32 (±5) nm between cytb_{562} molecules which is in line with the calculated distance between the anchoring points bound to every other DX-A tile (Figure 3.15(c)).[103,122] The pattern of the DX-array was not as visible compared to the images captured by liquid AFM, but in general, it is comparable with Rusling's AFM image of DX-array, biotin-TFO and an excess of streptavidin (SA), also captured with dry AFM (Figure 3.15(b)). The amount of protein recruited on the DX-array depends on the concentration of the added protein. Here, the amount of protein was not in excess, resulting to fewer molecules being attached to the DNA surface, whereas in Rusling's case, there was an excess of streptavidin. Considering all the above, the results demonstrate successful hybridisation of the protein to the recognition sequence of the DX-A tile.

3.2.6 CD & SRCD Spectroscopy of the DX tiles and DX-Array

The subsequent sections present the CD and SRCD experiment of the DX-array with or without protein attachment, which were recorded at B23 beamline at Diamond Light Source, during beamtimes allocations including SM18017-2, SM18017-3, SM19117-1, SM19117-2, SM20669-1 and SM20669-2. CD measurements in solution were recorded using the bench-top instrument, Chirascan Plus, whereas the SRCD and CDi experiments using films were performed in the module B station. Similar to Chapter 2, all the CD and SRCD data were processed in CDApps and Origin.

3.2.6.1 Thermal denaturation of cytb_{562} using circular dichroism

Prior to the CD experiments with the new DNA nanostructure (DX-array) and protein (cytb_{562}), the Chirascan Plus CD instrument was used to obtain information about the conformational change of the cytb_{562} associated with thermal denaturation. If the temperature at which protein denatures is known, it is easier to identify the conditions in which the protein handling can take place. The instrument was equipped with a Peltier cell (Quantum Northwest, Washington, USA) allowing adjustment of sample temperature. For the execution of this experiment, the LA cytb_{562}, after its modification with ssDNA, was chosen with a concentration of 3.5 μM in 1x phosphate buffered saline (PBS) pH 8.0. Conventional CD and UV-absorbance spectra of the protein were measured simultaneously in the near and far-UV regions at various temperatures. The measurements were performed in fused quartz cuvettes with 0.3 cm path length (Hellma, Essex, UK). The desired temperature range ranged between 20 and 90°C with programmed temperature

interval at 2.5°C. The temperature was raised using a stepped ramp with 2 minutes incubation at each point once the temperature had stabilised. After the last scan (90°C), a final scan at 20°C was recorded to monitor and quantify the protein re-folding after denaturation. Two consecutive repeated scans were collected for each temperature and then averaged, prior to subtraction of baseline spectrum of PBS buffer at 20°C. The CD and UV absorbance results of this experiment are shown in Figure 3.16.

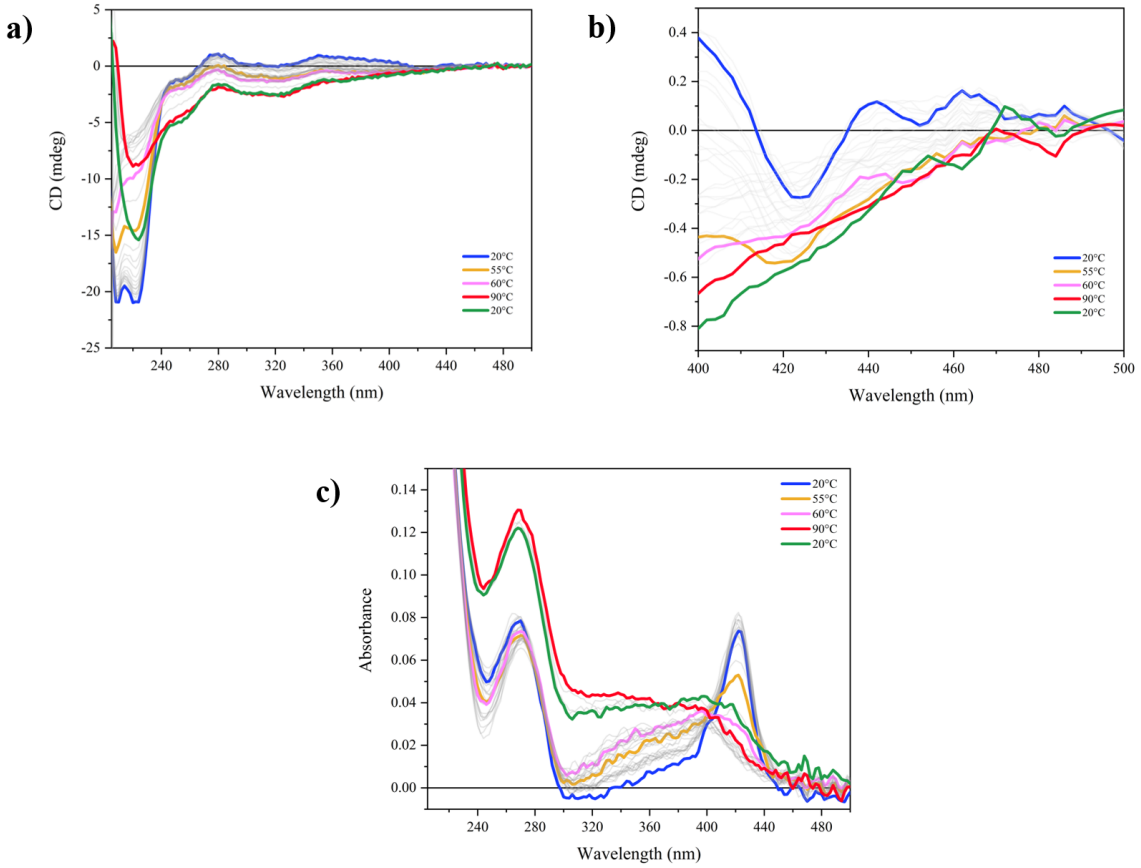

Figure 3.16: **a)** CD data of the LA cytb_{562}-ssDNA, full wavelength. **b)** CD data of the LA cytb_{562}-ssDNA, heme region and, **c)** absorbance data as they were all recorded during thermal denaturation experiment.

The temperature ramp experiment reveals that the protein under physiological conditions (starting at 20°C) adopted the expected helical structure for an α-helical protein characterised by the negative CD bands at ∼208 nm and at 220 nm (Figure 3.16(a)). However, the maxima at ∼195 nm is not visible due to cut-off. In addition, a negative peak at 420 nm was observed associated with the heme chromophore non covalently bound to the cytb_{562} protein (Figure 3.16(b)). The heme is an achiral molecule, but since it is bound to a chiral molecule, an induced negative CD band, the Soret band, at about 420nm is observed. The UV absorbance spectrum shows a sharp Soret band at around 420 nm, which is typical of a ferric heme (Figure 3.16(c)).[125] A sharp peak at 260 nm is also visible

which corresponds to the ssDNA that was incorporated to the protein through click chemistry. The ssDNA's contribution to the CD signature of the protein is represented with a weak positive peak at around 280 nm, characteristic for ssDNA.

As the protein is heated and begins to denature, there is a significant loss in spectral intensity, as expected, because the protein unfolds to become mere and more unordered. The double negative peak is constant over the first series of scans but in a lower magnitude, and after 35°C the CD signature of the minima alternates. By the time 60°C is reached, the entire shape of the curve is distorted since the protein is above the critical temperature, indicating a loss of structure. Similarly, the loss of the heme ligands with increasing temperature results in a noticeable shift of the Soret band in the visible spectrum, which is also stated in the literature.[131,132] This behaviour was expected, as heme-containing proteins (like cytochrome b_{562}) are particularly sensitive to the variations in temperature and, therefore, it is important to avoid excessive sample heating as they operate in a better way under physiological conditions.

Finally, in order to determine the exact melting temperature (T_m) of the LA cytb_{562}-ssDNA, the CD response obtained at 220 nm, which is one of the characteristic negative peaks of α-helical proteins, was plotted against the temperature and the melting temperature was determined by using the Boltzmann fit in Origin. According to the analysis, the (T_m) was calculated to be 55.5°C. Figure 3.17 presents the results of the Boltzmann fit analysis.

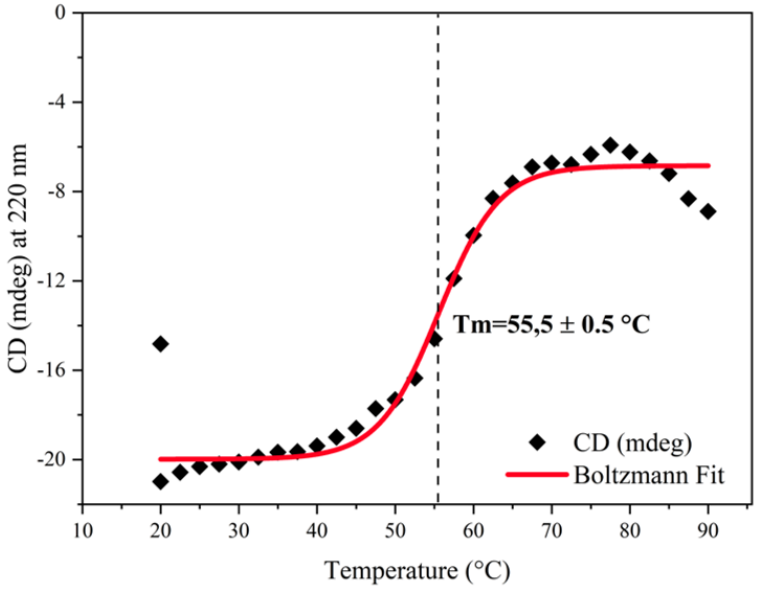

Figure 3.17: Melting temperature of the LA cytb_{562}-ssDNA at 55.5 °C as calculated using Boltzmann fit analysis in Origin.

The function of cytb_{562} is lost already at 40°C as indicated by the plateau from 20 to 40°C of Figure 3.17 and where the α-helical is fairly maintained. Above

40°C the heme starts to dissociate from the bound protein, which in turn starts to denature to an unordered structure as it is also evident from Figure 3.16. The process is not reversible at all with the protein back at 20°C adopting a mainly β-strand conformation and the achiral heme molecule remaining unbound. CD is very sensitive to chiral molecules and it is a very useful technique since it is the only one that can see only the bound species of achiral ligands. In the case of cytb_{562}, as the protein does not absorb in the Soret region in general, CD is an unambiguous method to determine ligand binding.

3.2.6.2 CD measurements in solution

CD measurements of the DX system before and after protein hybridisation were recorded in solution prior to the SRCD experiments to inspect the CD in terms of correct profile shape and intensity magnitude of this new DNA nanostructure system and in order to assess which procedure would give the best film sample preparation. Cylindrical cell made of fused silica and path length 0.02 cm was used for the DX-A tile and array, whereas a fused silica cuvette with path length 0.3 cm was used for the protein samples before and after attachment to the DX-A tile and array. 30 μL of each sample were pipetted into the cell or cuvette and the average of three repeated consecutive CD and UV absorbance scans were recorded at 20°C. The UV absorbance data of the DX-A tile and array were normalised in order to be comparable with the rest of the samples since a different cuvette path length was used for their recordings. It is also important to note that all samples were in 1 μM concentration except for the LA cytb_{562}-ssDNA and SA cytb_{562}-ssDNA which had 7 μM and 4 μM concentration, respectively. The protein/DNA ratio for the DX system with protein was 1:1, with each component's concentration of 1 μM. The data collected during this experiment are presented in Figure 3.18.

The CD spectral features of the DX-A and DX-array are weak, however the characteristic positive band at ∼280 nm and the negative one at ∼250 nm are noticeable and consistent with the related UV band at ∼260 nm. Also, the characteristic negative double CD bands at 222 and 208 nm associated to the protein helical conformation, though small in intensity, could be seen for both LA cytb_{562}-ssDNA and SA cytb_{562}-ssDNA. The positive band of the DNA included in the two variants is also present, along with the negative peak for the heme at 420 nm, which can be seen in more detail in Figure 3.18(b). The CD Soret band is an indication that the folding of the cytb_{562} is retained when attached to the DNA as demonstrated in the thermal denaturation experiment.

When the proteins are hybridised to the tile (DX-A with LA cytb_{562}, the DX-A+SA cytb_{562} data are missing), the negative CD Soret band is very small but it

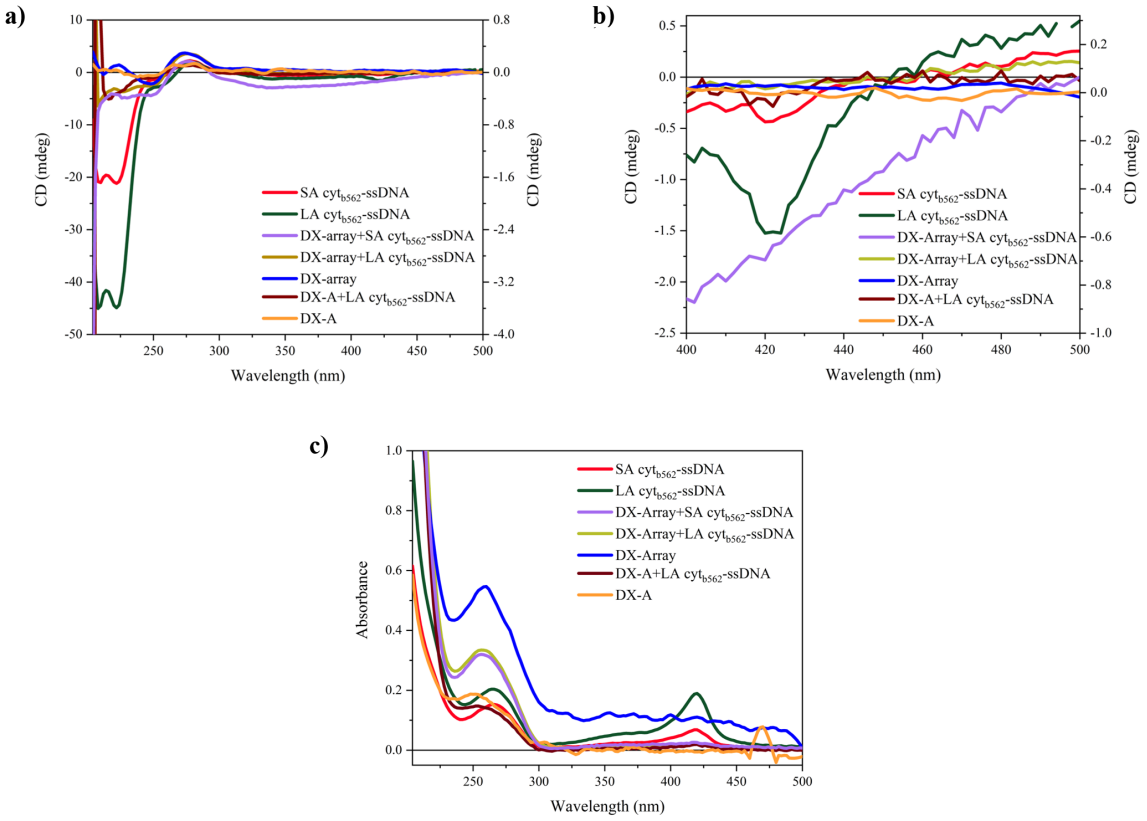

Figure 3.18: Conventional CD **(a, b)** and UV absorbance data **(c)** of the DX-A tile, DX-array and of the different variants of cytb_{562} before and after their hybridisation on the tile or array as captured in solution.

is there and since this negative CD band is an unambiguous indication that the heme is non covalently attached to the protein, it means that the native folding of cytb_{562} is retained as otherwise the CD Soret band would vanish as demonstrated in Figure 3.16 and 3.17. However, when the protein is attached to the DX-array, the negative band at 420 nm is not detectable or very weak. Similar behaviour is observed with the UV-absorbance spectra. This experiment was repeated by examining the different variants of cytb_{562} before and after the attachment to the DNA nanostructure all at the same concentration, by using different cuvette path lengths and by investigating the DNA-protein and the heme regions separately. The different path length could help increasing the CD signal of the DX-array and DX-A tile with the attached protein.

It is important though to keep in mind that the DX-A+protein and the DX-array+protein systems have not been purified any further after their hybridisation. That means that it is likely to have unbound protein in the samples that can easily influence the final data. Especially in the case that the data of the DX-array are subtracted from the DX-array+protein, the resulting spectrum may not reflect only the protein that it was bound to the array but also the free one. This can give mixed information about how and if the protein structure is affected after

its hybridisation to the DNA nanostructure. Unfortunately, until a way to purify the system is found, there is anything else that can be done in order to avoid this phenomenon.

Q71 cytb_{562} in solution

Initially, the Q71 cytb_{562} (short axis, monomer) was investigated in solution, both before and after its hybridisation to the DNA nanostructure. The control samples used were the wild-type holocytb_{562}, the Q71-cytb_{562} with the azF or ssDNA modification, along with the DX-A tile and the DX-array. Considering the experiment in solution described earlier, the induced CD of the heme is weak and as a result a 3 mm path length cuvette is not ideal for exploring this region (350-500 nm). Therefore, a 1 cm path length was finally selected. In addition, a 0.5 mm path length cuvette was firstly used for the protein/DNA region (200-350 nm) but the signal of the DNA was not strong enough. The absorbance was ∼0.1, which would result into 0.05 after the buffer subtraction. Even when the protein was on the DNA, the absorbance would not be more than 0.2 (the data of this recording are not presented). Consequently, a 3.0 mm path length was finally employed for the investigation of the protein/DNA region. Three scans were recorded for each sample and reported in the following figures.

In Figure 3.19, the controls DX-A tile and DX-array show the expected DNA CD signature with positive bands at 275 nm and 220 nm and a negative one at around 245 nm. As expected, the CD intensity of the DX-array is higher than that of the single tile's signal. On the other hand, the wild-type holocytb_{562} appears to have slightly lower α-helicity (lower magnitude) compared to the Q71 cytb_{562}-azF or Q71 cytb_{562}-ssDNA, indicating that the secondary structural changed upon azF and ssDNA binding. As Mart *et al.* reported in their research, the attachment of ssDNA impacts the activity.[24] The secondary structure of cytb_{562} appears to change after the addition of other compounds as it was also noticed when the holocytb_{562} was compared with the apoprotein in Kamiya's research.[132] The latter yielded lower intensity, so reduced α-helicity, compared to the former, indicating change in the secondary structure when the heme was incorporated.

The attachment of the azF or the ssDNA does not appear to affect the CD Soret band of the heme though (Figure 3.20), since the induced CD for all the control protein samples has similar intensity, which is consistent with the corresponding UV spectra. Similar behaviour is observed for the heme even after the hybridisation of the protein to the DX tile and array as the shape and the CD intensity are similar among all the samples. Therefore, from the CD of the Soret band, the protein is qualitatively intact and should retained its function related to the heme. However, the CD of the protein when it is immobilised to the DNA nanostructure

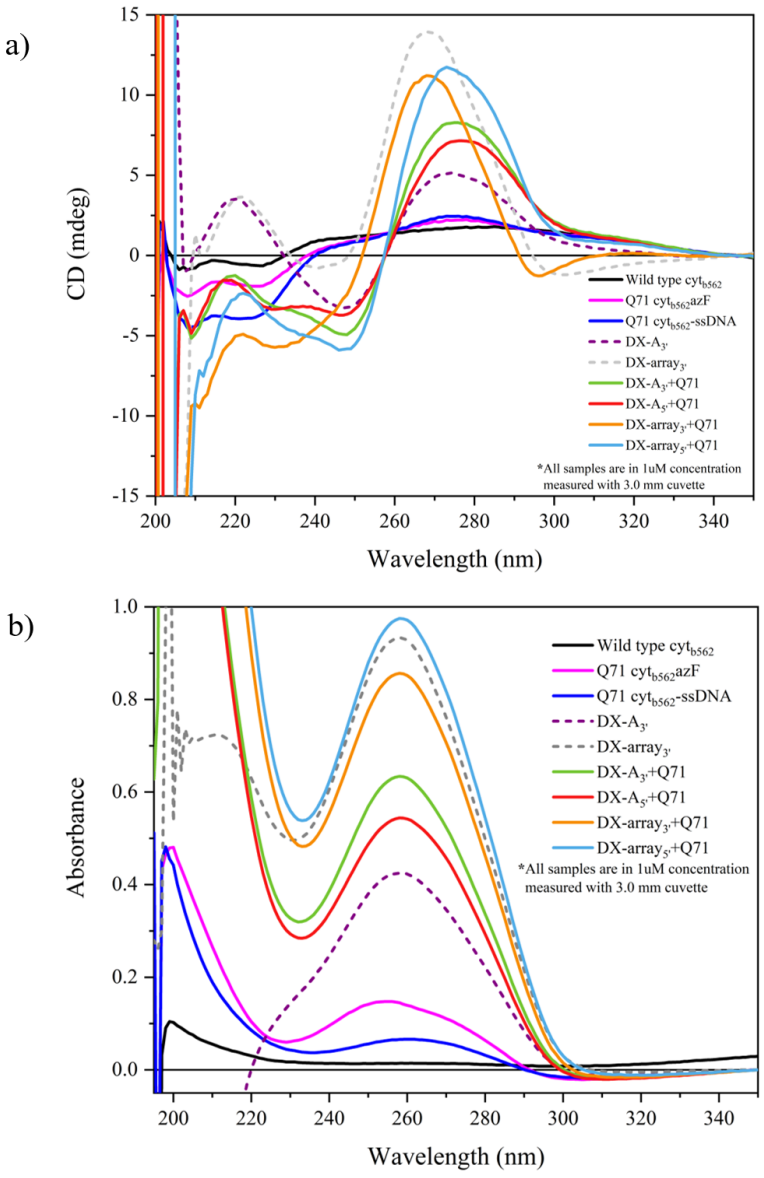

Figure 3.19: Conventional CD **(a)** and UV absorbance **(b)** data of the Q71 cytb_{562} before and after hybridisation to the DNA nanostructure. The protein/DNA region (200-350 nm) was scanned with 3.0 mm path length cuvette.

(DX-A$_{3',\text{ or }5'}$+Q71 or DX-array$_{3',\text{ or }5'}$+Q71) is affected and the expected minima at 220 and 208 nm are not detected. The shape of these CD spectra features resemble the CD signal of the DNA control samples rather than the protein control samples. It is expected that the CD spectra of the DNA are predominant between 250 and 300 nm as it is a stronger chromophore and the contribution of the protein's aromatic chromophores is weaker.[133] Nevertheless, it was also predicted that the two negative characteristic peaks of the protein would be present, which unfortunately is not the case.

In Figure 3.21, the CD data of Q71 cytb_{562} are pesented as they are configured after the subtraction of the DX-A$_{3',\text{ or }5'}$ or DX-array$_{3',\text{ or }5'}$ from their respective

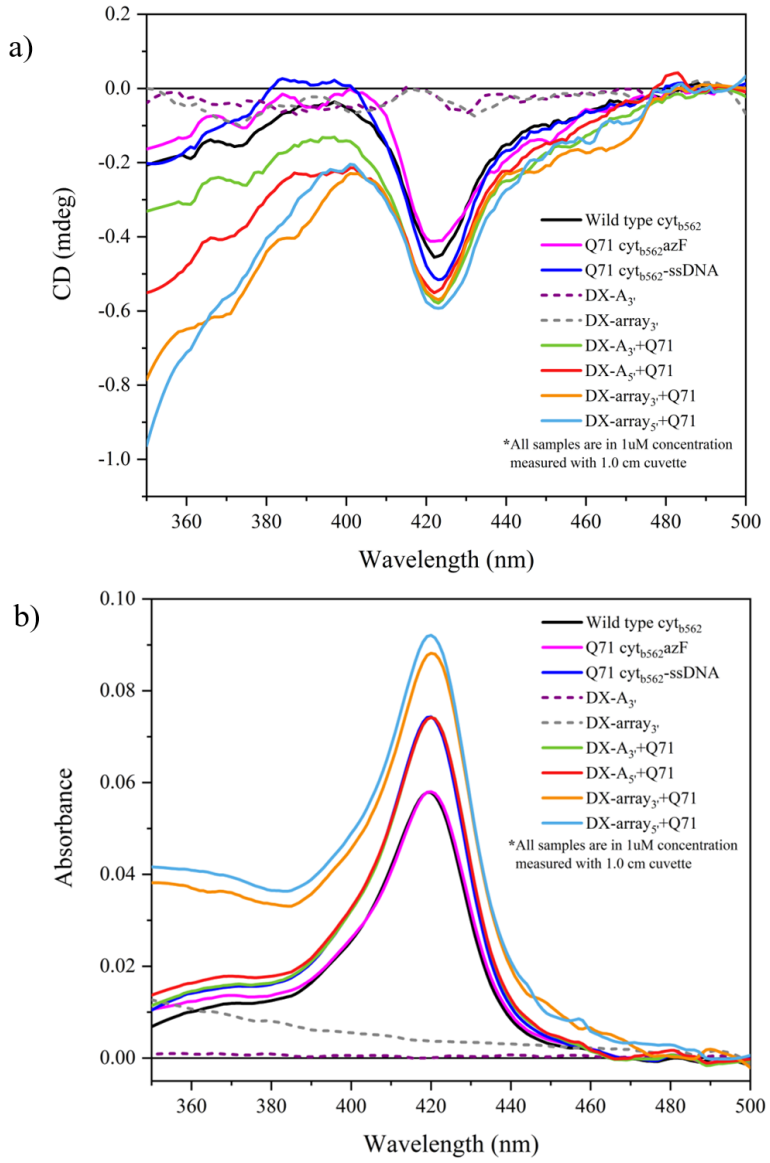

Figure 3.20: Conventional CD **(a)** and UV absorbance **(b)** data of the Q71 cytb_{562} before and after hybridisation to the DNA nanostructure. The heme region (350-500 nm) was scanned with 1.0 cm length cuvette.

samples with Q71 attached. This subtraction was carried out in order to examine the concept mentioned earlier by removing the DNA CD signal and ideally getting the CD signature just for the protein. As predicted, the signal at the heme region is not affected and remains exactly the same (Figure 3.21(b)). However, at the protein/DNA region, it seems that the CD signal for Q71 cytb_{562} (red and green solid line) resulted from the subtraction of DX-A$_{3'\text{ or }5'}$ from DX-A$_{3'\text{ or }5'}$+Q71 has changed in comparison to the one from Figure 3.19 revealing the expected protein CD signature with the minima at ∼220 and 208 nm. It is notable that the position of the protein on the tile, either positioned closely to the tile surface (DX-A$_{5'}$) or protruding away from it (DX-A$_{3'}$), does not appear to affect the signal significantly,

giving a similar result in both cases, which are also comparable to the Q71 cytb_{562}-ssDNA (purple fainted line), as expected. Only the negative peak at 208 nm from the Q71-ssDNA with CD signature indicated with the red line looks smaller than usual. This might be due to the original DX-A+Q71 sample (before subtraction) and the fact that the readings below 210 nm were slightly affected by the cut off.

In principle, when the structure of the protein, or molecule in general, after its hybridisation with other molecules does not change, the spectra should be additive. In this case here, it seems that when the protein is hybridised to the tile, the spectrum changes leading to the speculation if this is influence of secondary structure of protein or some other impact of exciton coupling.

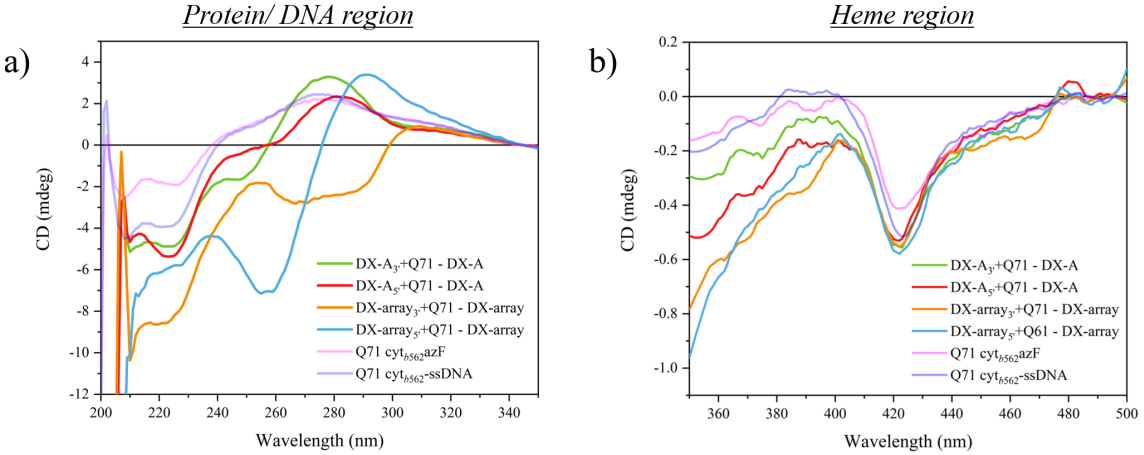

Figure 3.21: CD data of the Q71 cytb_{562} after the DX-A$_{3'\text{ or }5'}$ or DX-array$_{3'\text{ or }5'}$ subtraction as recorded at the protein/DNA region (a) and heme region (b). The solid fainted lines belong to the CD signal of the Q71 cytb_{562}azF (pink) and Q71 cytb_{562}-ssDNA (purple) used for comparison.

When it comes to the subtraction of the DX-array$_{3'\text{ or }5'}$ (Q71-ssDNA blue and orange line), the characteristic peaks are not visible and in general the CD signatures do not feature the signal of neither the DNA nor the protein. In addition, the two signatures behave quite differently from each other, which means that the position of the protein when it is immobilised on the bigger DNA nanostructure (closer or far from the array surface) may play an important role. Moreover, the CD signature might differ from the typical one of an α-helical protein because there could be a change in the secondary structure upon hybridisation. The most probable scenario is that since the readings of the original samples (not subtracted ones, Figure 3.19) around 210 nm are affected by the buffer which absorbs around this region, important information is missing after the subtraction. Therefore, the SRCD experiments of these samples will shed light on those scenarios. The CD and UV measurements using the other cytb_{562} variant should indicate whether the observations mentioned in this section are due to the protein.

Q50 cytb_{562} in solution

Q50 cytb_{562} (long axis, dimer) was the last protein investigated in the series of the experiments in solution before and after its hybridisation to the DX tile and array. The wild-type holocytb_{562}, the Q50 cytb_{562} with azF and ssDNA modification were used as control samples, as well as the DX-A tile and DX-array. The 1 cm and 3 mm path length cuvettes were chosen for the heme and protein/DNA region respectively. As described in in the Q71 cytb_{562} experiment, three scans were recorded for every sample and only the average was plotted upon subtraction against the respective buffer. All samples were in 1 μM concentration in order to be comparable.

In Figures 3.22 and 3.23, the conventional CD and UV absorbance data are plotted for the protein/DNA and heme region separately. The data displayed for the reference samples DX-A tile and DX-array were discussed in detail in the experiment of the Q71 cytb_{562} experiment and, therefore, they are not analysed further.

Similar to Q71 cytb_{562}, Q50 cytb_{562} with either the azF or the ssDNA modification show higher magnitude than the wild type holoprotein and between these modified proteins, Q50 cytb_{562}-ssDNA gives greater signal for the minima at 220 and 208 nm. However, when it comes to compare the two different variants of cytb_{562} (Q71 and Q50, short and long axis respectively) with the ssDNA, it is evident that Q50 yields greater signal ($\sim$ -8 mdeg), double than Q71 ($\sim$ -4 mdeg, Figure 3.19(a)) at 220 and 208 nm. According to this result, it is evident that the position where the ssDNA is incorporated into the protein, either in the short or long axis of the protein, plays an important role to their CD signature.

Furthermore, the heme region shows changes between the reference samples of the Q50 cytb_{562} variant. As mentioned earlier, the literature states that cytb562 tolerates azF incorporation, which was also confirmed in the case of the Q71 variant, even when the modification occurs in the residue 50 which is very close to the heme cofactor.[110] Nevertheless, the absorbance (Figure 3.23(b)) of the Q50 cytb_{562}-azF does not exhibit this behaviour since the absorbance is greater than the wild-type, but it could be a mistake in the stock concentration and the final concentration derived from the dilution of the sample. In addition, when the ssDNA is introduced in the Q50 variant, it yields stronger CD minimum at $\sim$ 420 nm than the wild-type (Figure 3.23). Compared to the respective sample for the Q71 variant, it is evident that the CD is enhanced for the Q50 variant. Hence, the incorporation of the ssDNA in residues closer to the heme cofactor seems to boost the CD signal.

When the protein is hybridised to the DX-A tile and DX-array (DX-A$_{3'\text{ or }5'}$+Q50 or DX-array$_{3'\text{ or }5'}$+Q50), the same behaviour as for the Q71 variant is observed.

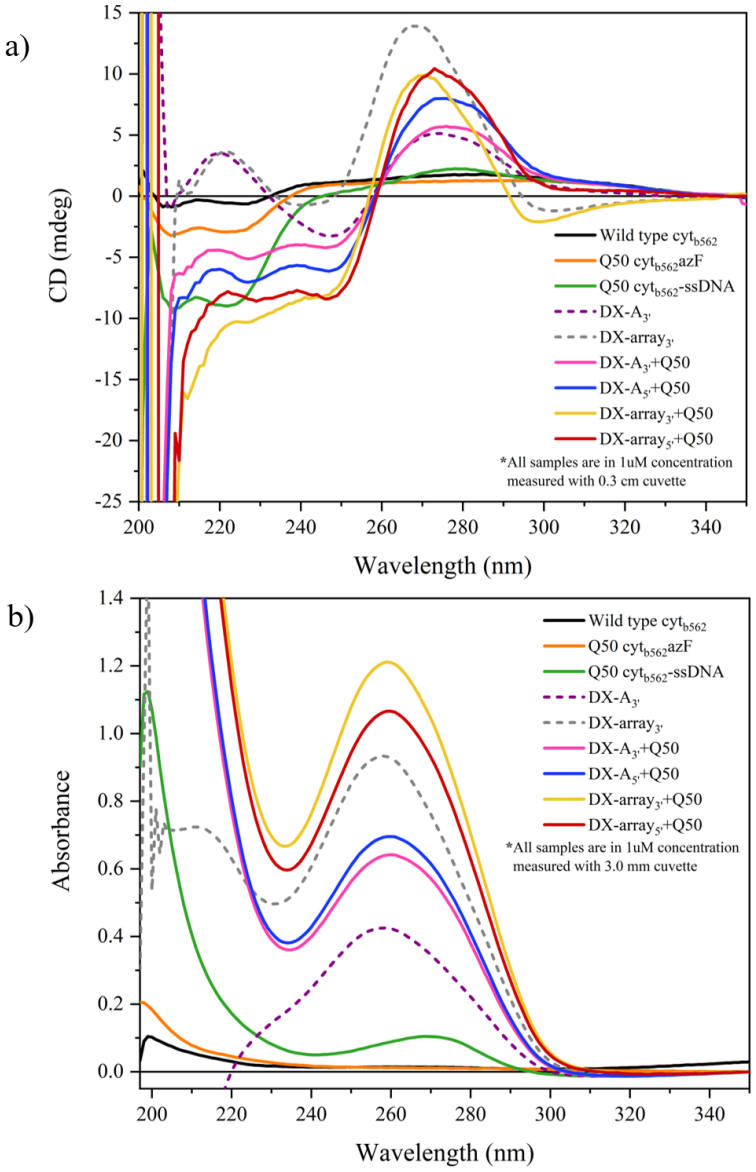

Figure 3.22: Conventional CD **(a)** and UV absorbance **(b)** data of the Q50 cytb_{562} before and after hybridisation to the DNA nanostructure. The protein/DNA region (200-350 nm) was scanned with 3.0 mm path length cuvette.

The positive peak at 270 nm for the DNA is present in all samples but expected minima are not visible, most likely due to the fact that the CD signal of the DNA is dominant, since it is a stronger chromophore and the protein is not in excess. In the case of Q50 variant, having the protein close or far from the array obviously has an impact to the spectrum, where the protein signals are somewhat weaker in the DX-array$_{5'}$ compared to the DX-array$_{3'}$. Furthermore, the DNA signal is also affected and the maximum at 275 nm becomes broader when the protein is close (DX-array$_{5'}$). Although the signal below 240 nm does not match accurately the protein's signal, compared to Q71, the Q50 variant shows greater intensity in this wavelength region than Q71, verifying the results discussed earlier. Therefore, the

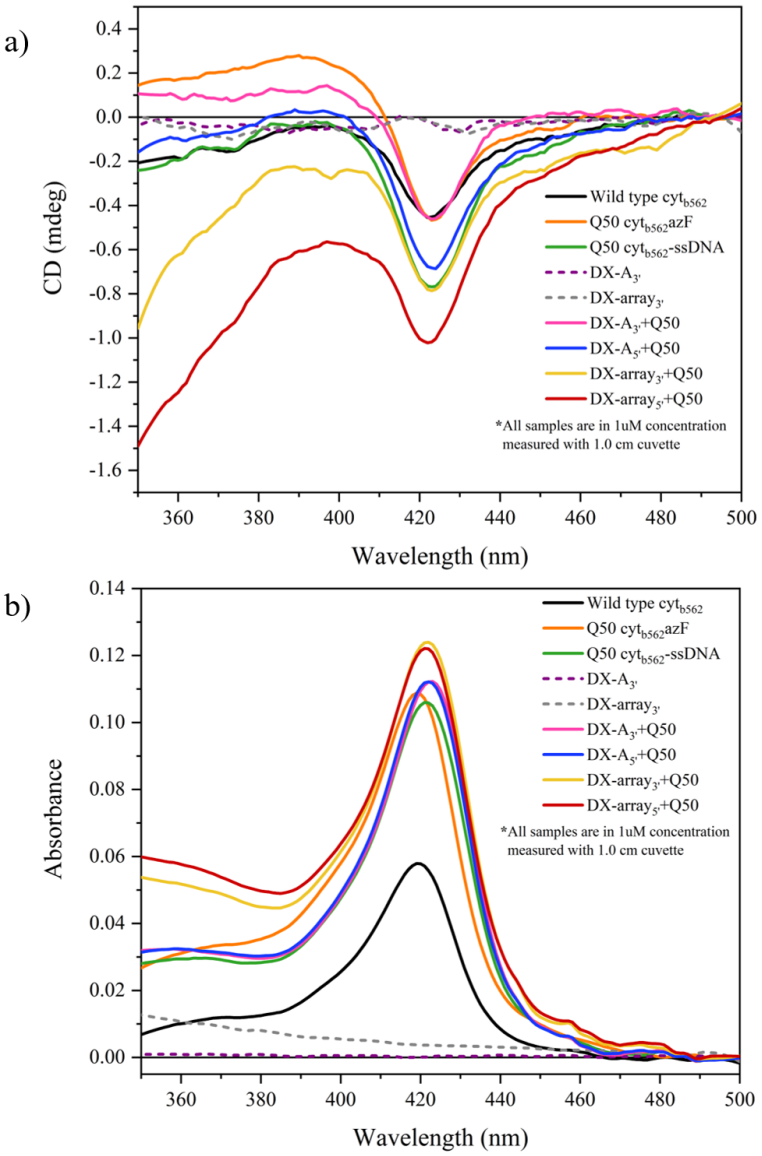

Figure 3.23: Conventional CD **(a)** and UV absorbance **(b)** data of the Q50 cytb_{562} before and after hybridisation to the DNA nanostructure. The heme region (350-500 nm) was scanned with 1.0 cm length cuvette.

long axis (Q50) variant, which is attached to the DNA nanostructure vertically, increases the CD signal in contrast to the short axis variant (Q71) which is attached horizontally, and therefore in the proximity of the DX-array surface which may lead to influence on their structure and decrease in the overall signal. Taking all the above into account, this method can be used to distinguish protein variants when they are immobilised on surfaces and their position is not known. However, in many cases, the signal intensity is also influenced but the alignment of the protein to the beam, e.g. perpendicular or parallel.

The subtraction of DX-A$_{3',\text{ or } 5'}$ or DX-array$_{3',\text{ or } 5'}$ from their respective samples with the Q50 variant attached was carried out similar to the Q71 variant, in order

to isolate the protein CD signal by removing the DNA. The CD data resulted by this subtraction are presented in Figure 3.24. The heme region is not affected by the DNA as it was expected since the DNA does not absorb in this region. On the other hand, in the protein/DNA region, the CD spectra for the protein samples where the protein was attached either to the DX-A tile or the DX-array change significantly. After the subtraction of the DX-$A_{3'\text{ or }5'}$, both CD signatures (magenta and blue, respectively) yield a minimum at 220 nm and almost at the same mdeg value that the reference sample was recorder (light blue), but unfortunately the second negative peak at 208 nm is not visible due to the buffer interference. When DX-$array_{3'\text{ or }5'}$ is subtracted, similar results to the Q71 variant are obtained, which did not show any characteristic peak of the α-helical protein or the DNA. Due to buffer absorbance at ~210 nm, significant amount of information is missing and the same challenges faced with the Q71 variant are also observed here. Therefore, the SRCD experiments are needed in order to collect the information that could not be collected from series of experiments.

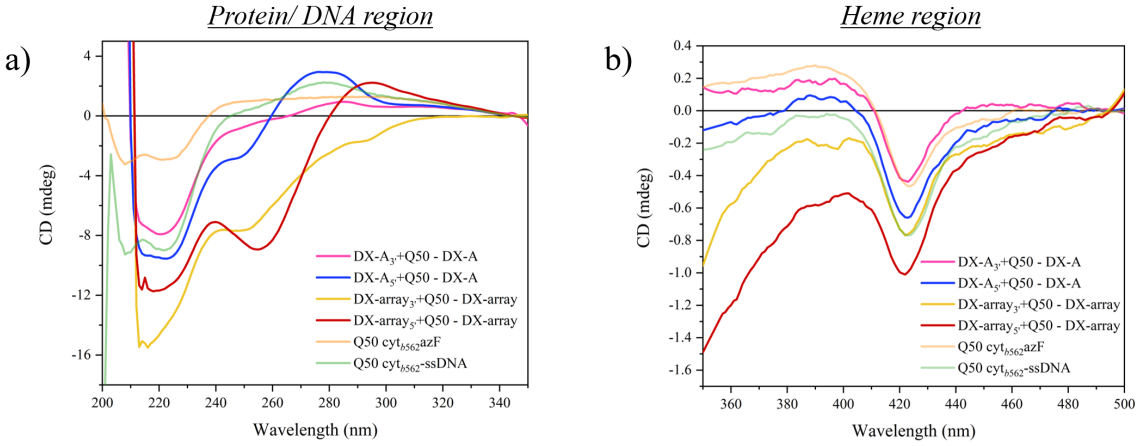

Figure 3.24: CD data of the Q50 cytb_{562} after the DX-$A_{3'\text{ or }5'}$ or DX-$array_{3'\text{ or }5'}$ subtraction as recorded at the protein/DNA region (a) and heme region (b). The solid fainted lines belong to the CD signal of the Q50 cytb_{562}azF (orange) and Q50 cytb_{562}-ssDNA (green) used for comparison.

3.2.6.3 CD experiments using films/ CD *imaging*

Thin films for the DX-A tile and the different cytb_{562} variants reference samples, as well as for the DX-array with and without protein attached were prepared to be studied under dry conditions using SRCD spectroscopy. Although the aim was to test the homogeneity of the films in terms of concentration and structure of all the films using the CD*i* method (Subsection 2.2.7.3), single scans were carried out for some samples, due to (a) either time constraints during the beamtimes or (b) poor quality of the films discouraging their use for further analysis with CD*i*. All films were prepared by depositing each sample at the centre of fused silica crucibles

and letting them dry under a blanket of nitrogen (Section 6.14). CD of the fused silica crucibles were used as baseline to be subtracted from the corresponding CD spectra with the deposited DNA-protein arrays.

DX-A tile, single scans

The DX-A tile was the first reference sample studied using SRCD spectroscopy. 10 μL of 1 μM sample were deposited on the fused silica substrate in order to form a thin film as described previously. In Chapter 2, it was concluded that it is very challenging to create uniform films which are homogeneous in terms of supramolecular structure with the appropriate thickness. In the case of the DX-A tile, it was noticed that when the 10 μL of sample were deposited on the substrate, the drying of the solution drop led to a non-uniform film. After visual inspection, a coffee ring effect was observed, resulting most of the sample to be accumulated at the edges of the film, leaving the film's central area much thinner than the border. Therefore, this film preparation was not appropriate for CDi and it was only used for single scan spectra in order to assess the supramolecular structure of the DX-A tile. Two different spots were selected to be scanned (Figure 3.25).

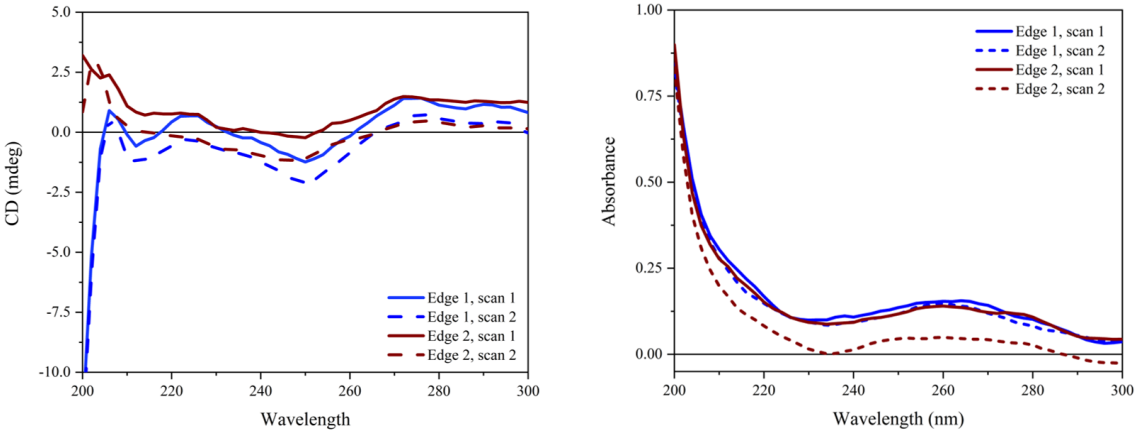

Figure 3.25: SRCD (left) and UV absorbance (right)) data of the DX-A tile (1 μM). Two scans (scan 1, solid and scan 2 dashed) were recorded for each spot of the film

The CD of the film of DX-A tile is weak in general, however a small broad positive peak at ~270 nm, a negative peak at ~250 nm and a crossover at 258 nm can be distinguished, which are similar to those observed in solution. Under other circumstances, it would be ideal to increase the signal of the sample in order to ensure that we record above the noise, but it was expected that the CD of the tile itself would be weak since the same was also observed for the CTB tile of the 2D DNA nanogrid in Chapter 2.

DX-array, CD *imaging*

Thin film of the DX-array was prepared by depositing 15 μL of sample in quartz substrate and drying it slowly under a stream of nitrogen. The film was visually

inspected appearing fairly homogeneous in terms of thickness, and most importantly devoid of any coffee ring effect. The DX-array film area was scanned as a 4x4 grid, measuring 16 spots at the spatial interval of 500 μm. The absorbance spectra were recorded simultaneously and the system was calibrated in both cases with regard to the background spectrum of the blank fused silica substrate as CD baseline. The 16 CDi spectra measured for the DX-array film (Figure 3.26(a)) were analysed by processing the data into two distinct 2D colour maps, which were generated from the intensity at fixed wavelengths at 275 nm and 220 nm for the 4 x 4 grid arrays (Figure 3.26(b), right and left respectively), as well as the corresponding 3D colour maps. Each square represents the spot mapped.

In Figure 3.26(a), two main peaks are visible, a positive one at 275 nm, a negative shoulder at about 245nm, and a negative one at ∼220 nm. The negative band at 220nm has been observed, though shifted to 210nm for the AB film of preparation F1 (Figure 2.28). Interestingly, the sign of this band was positive for the AB film of F2 preparation (Figure 2.27). The same sign but shifted to 230nm was also observed for CTB tile (Figure 2.25). Since there is limited information about DNA nanostructures in films and their SRCD or CD spectral features, it is hard to speculate what the real reason causing this shift is. In general, the SRCD signal for the DX-array film is more intense in comparison with the one recorded for the 2D DNA nanogrid indicating the better quality of the DX-array film. The 16 SRCD spectra were qualitatively very similar with small intensity changes particularly for the 220nm negative band.

The variations in the intensity can be explained by examining the absorbance spectra in Figure 3.28(a) or in Figure 3.27(right) where the absorbance spectra are superimposed next to the CD for the corresponding area that was mapped. It is evident that there are no critical variations among the absorbance spectra since, according to the 3D colour map (Figure 3.28(b)), the film fairly homogeneous in terms of concentration. Hence, the different intensities of the SRCD signal at 220 nm could be due to conformational changes of the DNA upon dehydration.

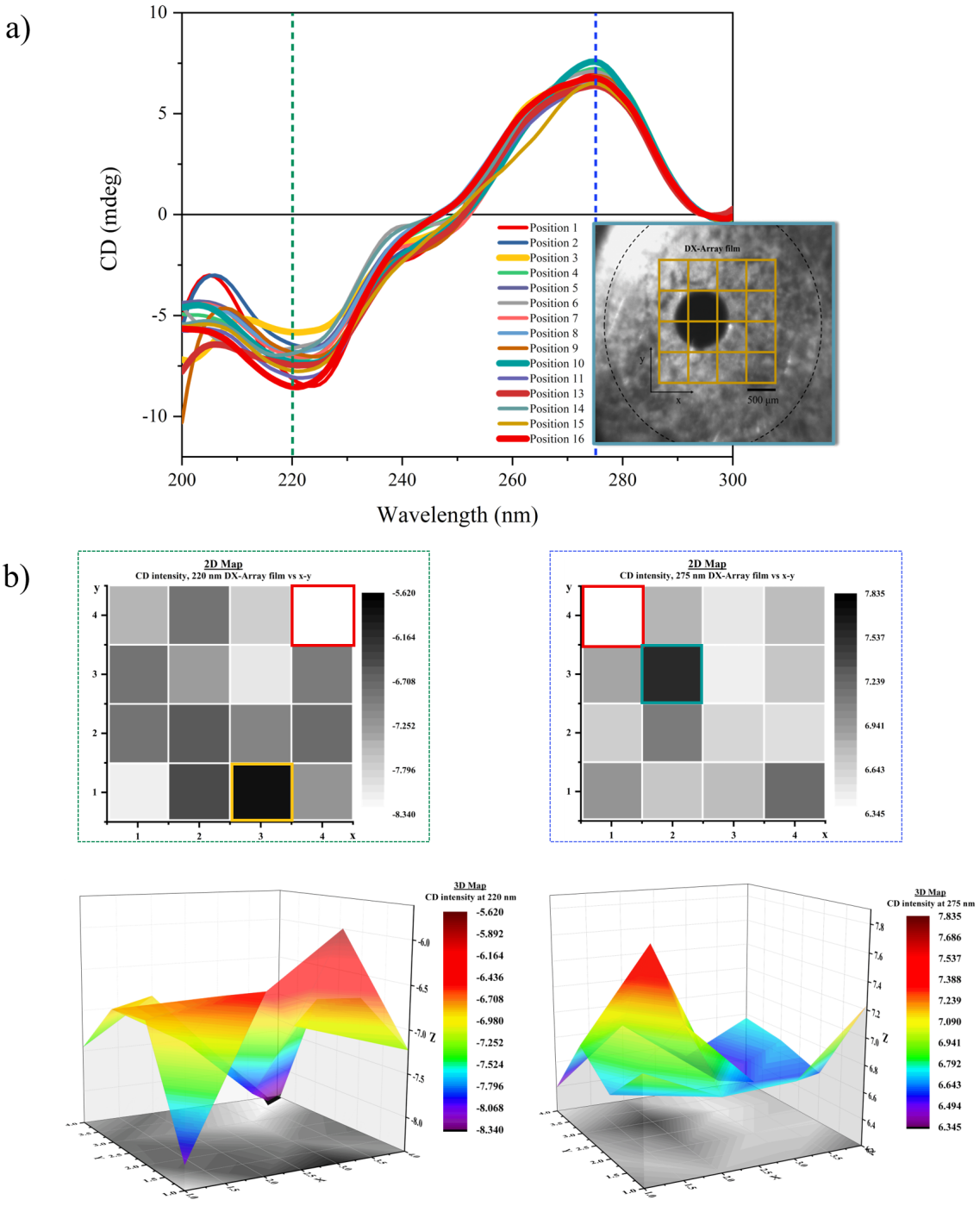

Figure 3.26: **a)** SRCD spectra of DX-array film after scanning a 4x4 grid array of 500 μm step size. An image of the film as captured with the camera of Linkam stage, showing the irradiated 4x4 sampling area. **b)** 2D and 3D colour maps at 220 and 275 nm (left and right, respectively).

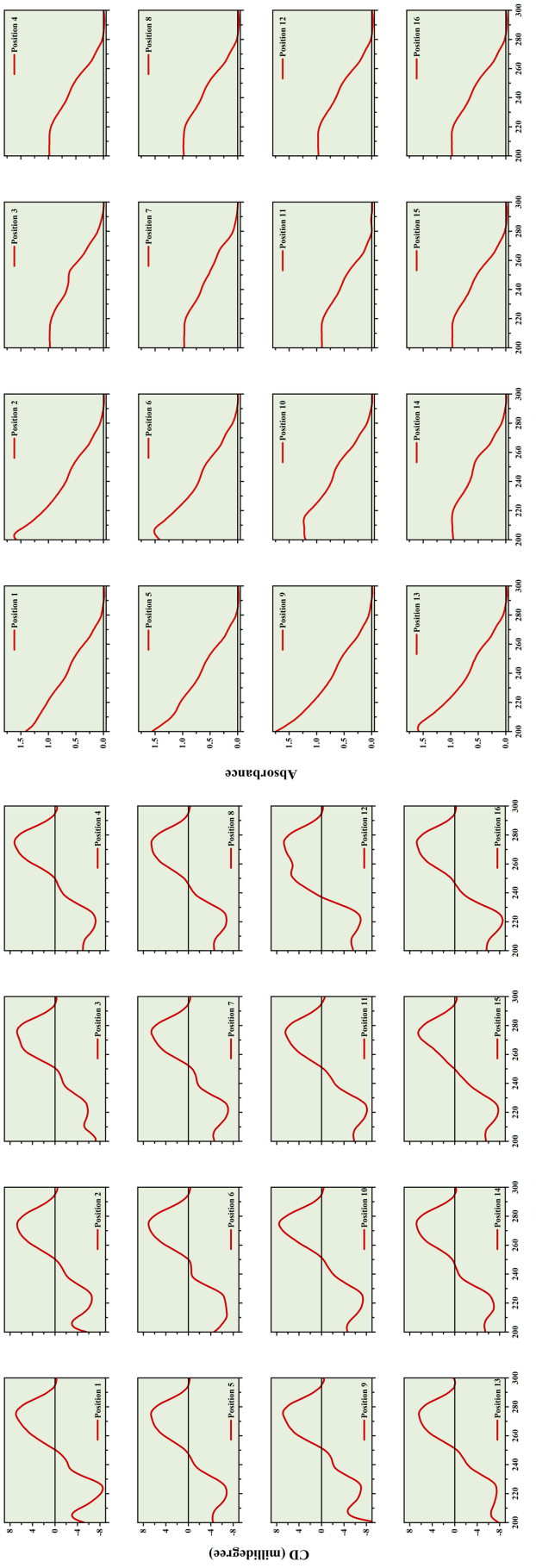

Figure 3.27: Square grid maps of SRCD spectra (left) and UV-absorbance spectra obtained after mapping a 4x4 grid array area with 500 μm step size.

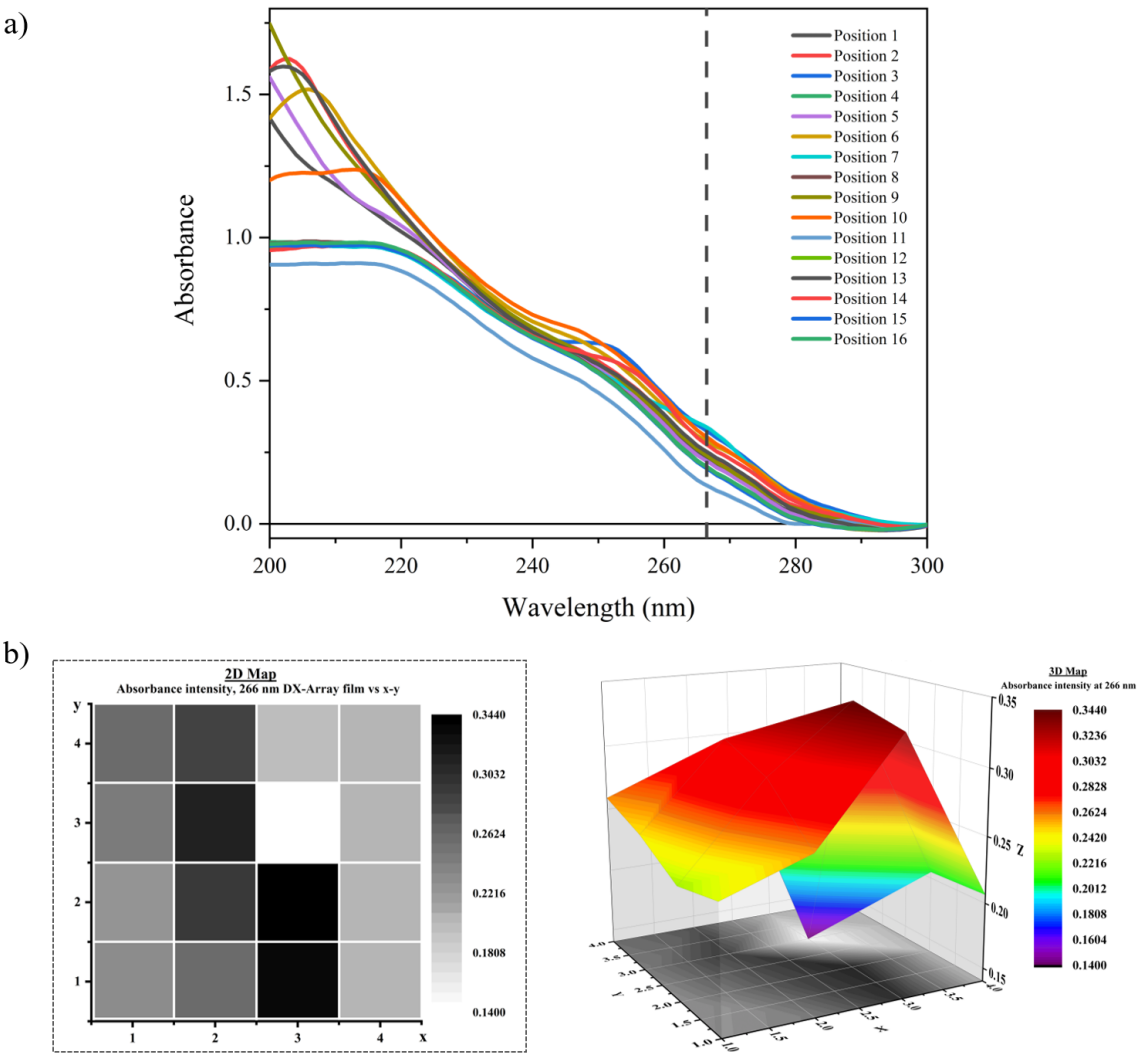

Figure 3.28: a) UV-absorbance spectra of DX-array film after scanning a 4x4 grid array of 500 μm step size. An image of the film as captured with the camera of Linkam stage, showing the irradiated 4x4 sampling area. **b)** 2D and 3D colour maps at 220 and 275 nm (left and right, respectively).

The DX-array film was further analysed by mapping a 50 x 50 grid array with 50 μm step size but at two distinct wavelengths. The wavelengths chosen were the one of the negative peak at 220 nm and the second one of the positive peak at 275 nm. For each experiment, 2500 scans were recorded and the data collected were processed into 3D colour maps and 2D image plots (Figure 3.29(a) and (b), respectively). Based on the 3D maps and image plots, it is deduced that the film is uniform. However, some scattered anomalies are observed in Figure 3.29(b), while the top-right corner of the plot shows greater inconsistency compared to the rest of the data (Figure 3.29(b), left). The film exhibits similar behaviour in Figure 3.29(a), even though the scattered anomalies are absent. This suggests that the beam hit the edge of the film, resulting into the deviation noticed at the top-right corner of the plots.

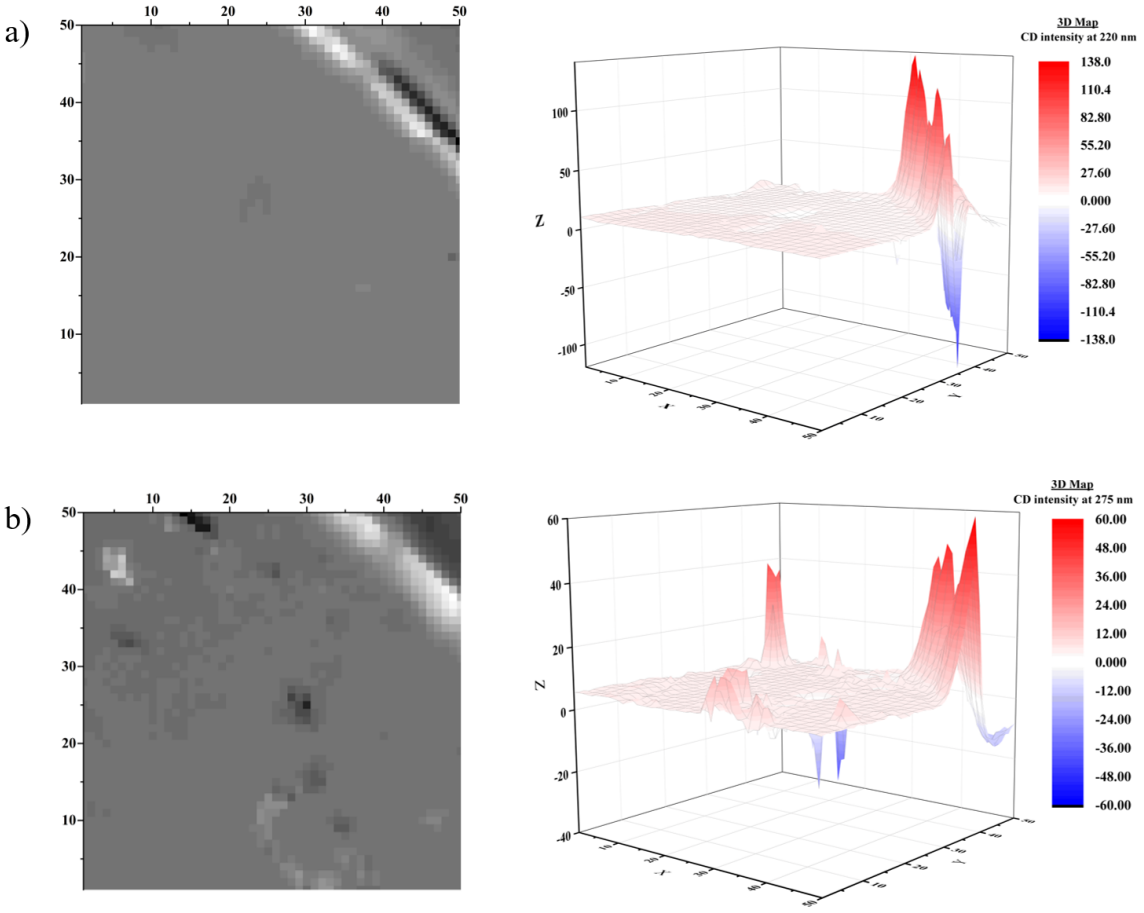

Figure 3.29: SRCD data. Image plots and 3D colour maps after single wavelength mapping of 50 x 50 points with 50 μm per step at **a)** 220nm and **b)** 275 nm.

Finally, the g-factor analysis introduced in Chapter 2 for the AB grid (Protocol F1) was also carried out for the DX-array film in order to investigate if there are any structural changes occurring when the DNA sample is prepared as a film (Figure 3.30). It is evident that the behaviour among the spectra in terms of intensity is divergent, leading to the conclusion that the film it is not homogeneous in terms of structure which also verifies the speculations made during the analysis of the SRCD data using the CD*i* method. In the literature it is stated that DNAs which belong to the B-form family can be forced to succumb to the A-form under dehydrating conditions.[74] However, it does not seem to apply in the case of the DX-array film since the characteristic peaks for a B-form DNA that goes into A form are even more positive CD at 270 nm with an intensive negative CD band at 210 nm. These do not appear to be the case as illustrated in Figure 3.26. However, the reduced positive CD at 275 nm could be consistent with dehydrated dsDNA.[74] An alternative potential root cause for this type of spectral change is that this reduction is linked to the interaction of neighbouring nucleotides and more specifically, is related to the increase in the winding angle and decrease in the base pair twist.[134] Even though, this information could explain the differences

in the CD signal and by extension in the structure of the DNA among the different areas mapped in the film, conclusions were made after analysing different kind of DNA samples than the DNA nanostructures. Since there are not many studies conducted about the CD of DNA nanostructures and especially as solid state films, it is not easy to draw any conclusions. Nevertheless, it is very likely that there are structural differences due to the base conformation upon its dehydration which may differ in each position mapped.

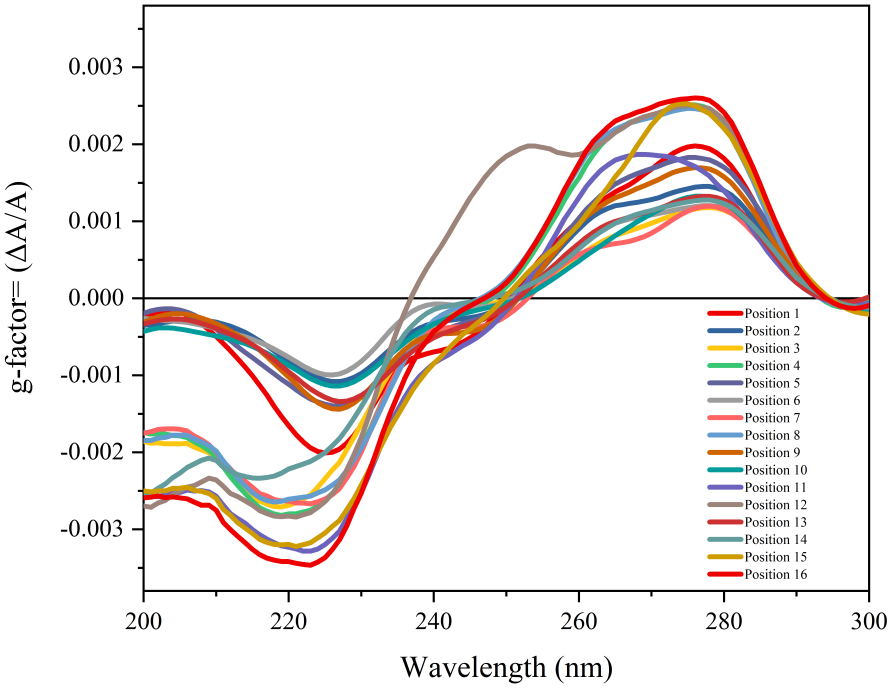

Figure 3.30: G-factor spectra for the DX-array film.

Cytb_{562} variants, single scans and CD *imaging*

The focused microbeam of B23 Beamline was further used to investigate the films of the different variants of cytb_{562} before and after their hybridisation to the DX-A tile and array. For each film created, 10-15 μL of sample were deposited and dried on the fused silica crucibles. SRCD and absorbance data were collected either during single scan at different positions on the films or during CDi experiments. The approach which would be followed for the measurements depended on the quality of the film after visual inspection. All data were calibrated with regard to the background spectrum of the blank fused silica substrate.

Initially, the dimers of the ssDNA modified SA (position 5) and LA (position 50) cytb_{562} were examined using single SRCD scans and the average of three consecutive repeated scans for each sample are shown in Figure 3.31. For each protein variant two sets of data are presented as they were recorded in different

beamtimes using different batches of protein. The samples used for the SA cytb_{562}-ssDNA were in 4 or 10 μM concentration (3.31(a), left and right, respectively) and for the LA cytb_{562}-ssDNA were in 7 and 10 μM concentration (3.31(b), left and right, respectively).

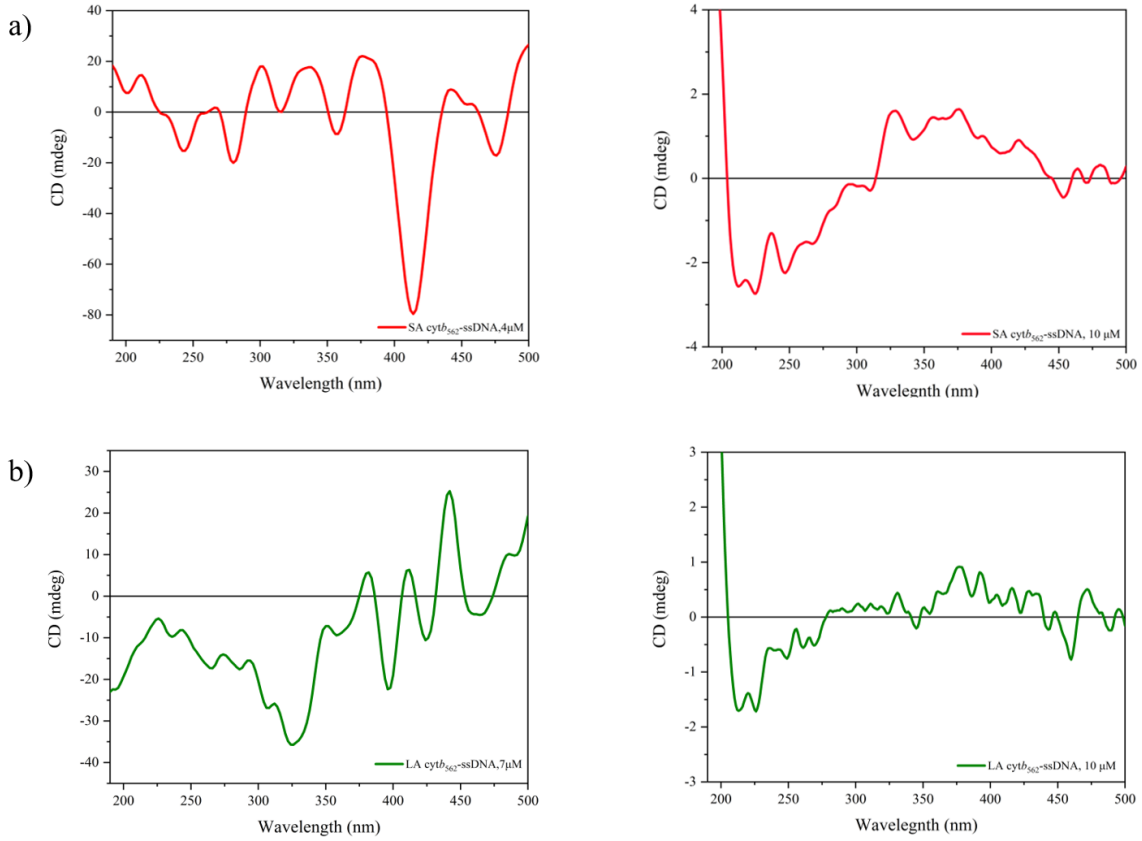

Figure 3.31: SRCD spectra showing the average of three scans for **a)** SA cytb_{562}-ssDNA at 4 (left) and 10 μM (right) and **b)** LA cytb_{562}-ssDNA at 7 (left) and 10 μM (right).

The results recorded for the SA and LA cytb_{562}-ssDNA at 4 μM and 7μM, respectively, are not in line with the data collected for the same protein variants (different batches and films) at 10 μM concentration. The spectra for the 10 μM protein samples resemble the spectra obtained in solution using the bench-top CD instrument (Figures 3.18, 3.19 & 3.22) yielding a characteristic minima at 220 nm and 208 nm. It is not surprising that the DNA peak is not visible since it was observed earlier that the DNA signal for the DX-A tile was very weak and, as a result, it would be very challenging to actually record the signal for an even smaller piece of DNA as ssDNA when the sample is being investigated as a film. As expected the induced CD signal for the heme is also absent since it showed low magnitude when the samples were investigated in solution, and it leads to the conclusion that the quality of the film was poor and therefore the data do not provide any useful information.

The following SRCD experiments are focused on the Q71 cytb_{562} variant (short axis, monomer). The samples tested were the azF and ssDNA modified Q71 cytb_{562}, along with the protein after its hybridisation either to the DX-A$_{(3')}$ or DX-A$_{(5')}$ (2:1 protein/DNA ratio). 10 μL of 1μM sample was deposited on individual crucibles in order to prepare the different films. After visual inspection, the coffee ring effect was noticed again, resulting the sample to be concentrated mostly at the edges of the film. Therefore, various positions of each film, mainly close to the edges, were probed and the results are presented in Figure 3.32. The solid lines correspond to the first scan of each position whereas the dashed and dotted lines of the second and third scans, respectively. The black, blue, red and green lines stand for the centre, the edge at the right-hand side (edgeR), the edge at the left-hand side (edgeL) and the edge at the bottom (edgeB) of the film, respectively.

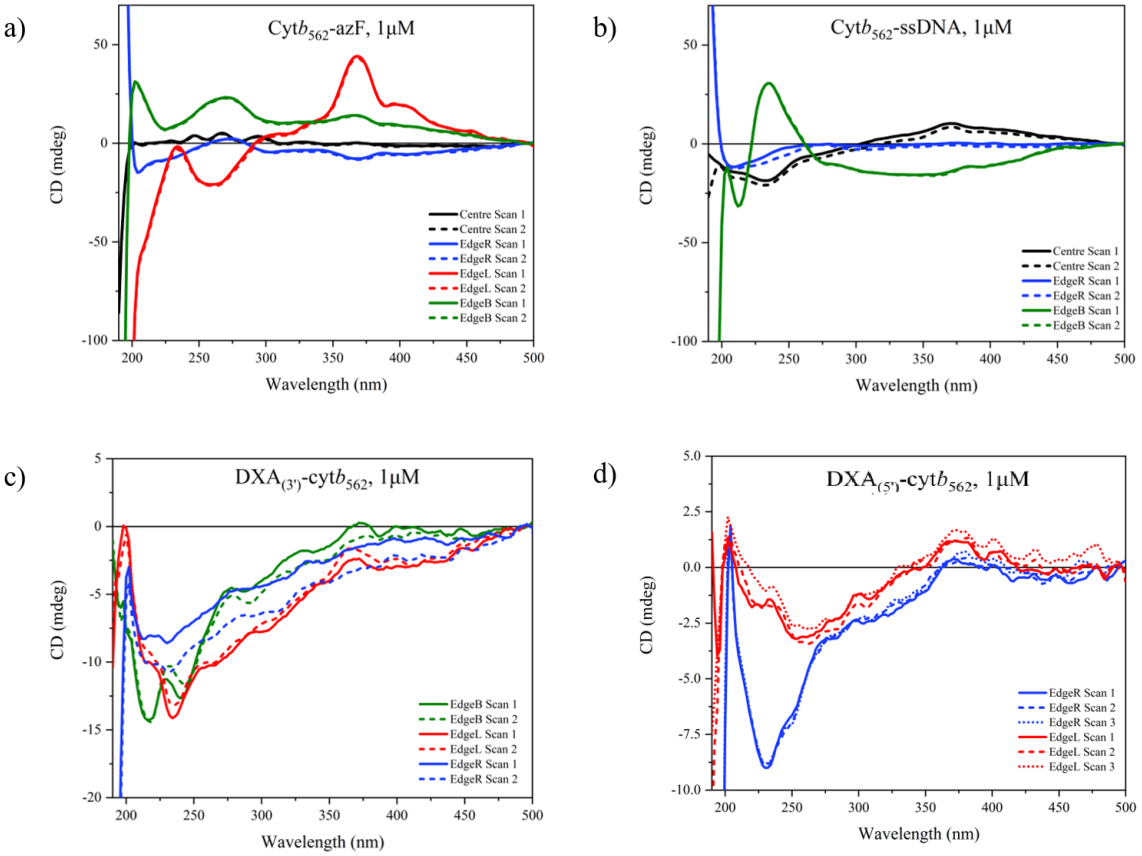

Figure 3.32: SRCD spectra for the films of the different Q71 cytb_{562} samples.**a)** Cytb_{562}-azF, **b)** Cytb_{562}-ssDNA, **c)** DX-A$_{(3')}$+cytb_{562} and **d)** DX-A$_{(5')}$+cytb_{562}. All samples are in 1 μM concentration and various positions of each film were probed.

The SRCD spectra show different behaviour from sample to sample, as well as among the positions of the same film. Some of the scans resemble the spectra of an α-helical protein. However, the negative CD associated to the Soret band is not detected or the signal is too weak when the protein is in solid phase as film. The Soret band should be the diagnostic signature and since it is not present here,

it could indicate that the sample is compromised, certainly in terms of the protein structure and therefore, it is hard to draw any conclusions from this set of spectra. In general, the signal is very weak which is reflected by the UV absorbance data in Figure 3.33. A general observation is that the films are not uniform in terms of structure and this does not seem to derive from the changes in the localised concentration (Figure 3.34) as it is fairly similar for almost all positions for each respective film.

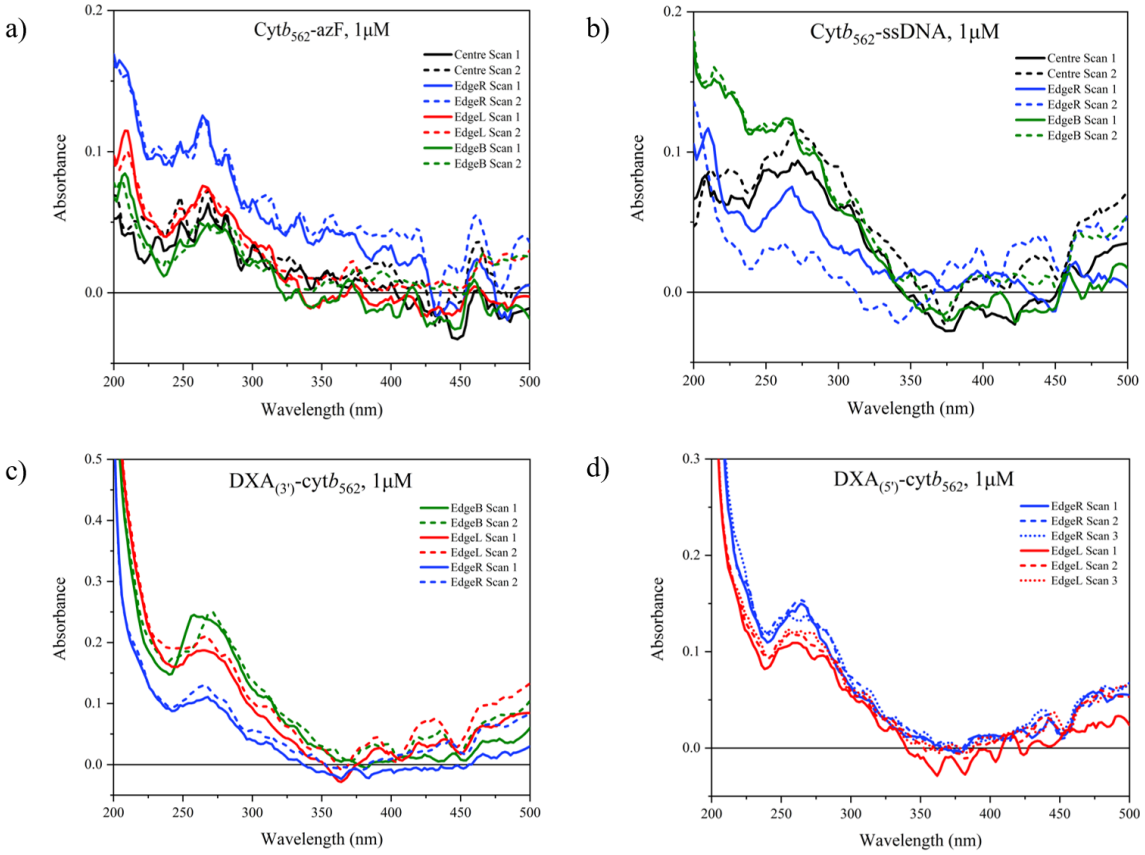

Figure 3.33: UV absorbance spectra for the films of the different Q71 cytb_{562} samples.**a)** Cytb_{562}-azF, **b)** Cytb_{562}-ssDNA, **c)** DX-A$_{(3')}$+cytb_{562} and **d)** DX-A$_{(5')}$+cytb_{562}.

Moreover, the Q71 variant was also hybridised to the DX-array (2:1 protein/DNA ratio) and the respective film was used for CD*i* experiment. A 3x3 grid array was scanned with 500 μm step size for the DX-array$_{(3')}$+cytb_{562} film, whereas a 10x10 grid array was probed for the DX-array$_{(5')}$+cytb_{562} but with smaller step of 150 μm. In Figure 3.34, the SRCD and UV absorbance spectra are presented, as well as the average of all the spectra recorded for each film. Here, the CD Soret band is also missing indicating that the sample is compromised. Some CD spectral features might give positive and negative bands characteristic for DNA or protein but it could be driven by luck.

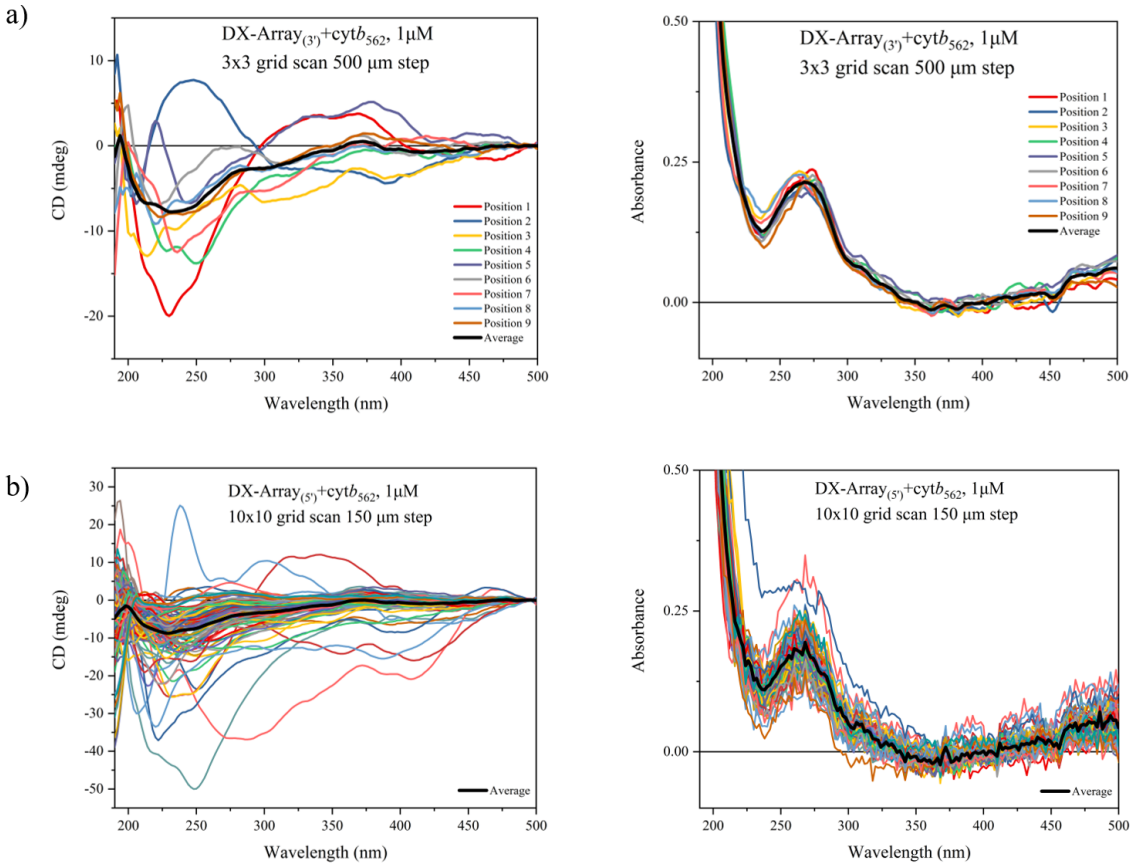

Figure 3.34: SRCD and UV absorbance data of **a)** DX-array$_{(3')}$+cytb_{562} film after scanning a 3x3 grid array of 500 μm step size and **b)** DX-array$_{(5')}$+cytb_{562} film after scanning a 10x10 grid array of 150 μm step size.

Finally, the diversity between the data in solution and in films is apparent and there is an obvious loss of intensity too, which derives from dehydration effect. In low hydration condition, proteins can aggregate which may lead to flattening effect or look more like crystallised which may result to scattering. In Figure 3.35, two different SRCD spectra of the DX-array$_{(3')}$+cytb_{562} are presented as they were recorded at the same position of the film. The only difference between those spectra is that the sample used in Figure 3.35(a) was still in its liquid form on the crucible and the sample in Figure 3.35(b) was half liquid, half dried. Two scans were recorded for each condition and the data were baseline-corrected against the clean quartz crucible.

It is clear that when the sample is hydrated, the two minima (210 and 220 nm) of the protein are present, as well as the positive characteristic peak for the DNA at 275 nm. As soon as the sample is dehydrated completely, as shown in Figure 3.34, or to some extent (Figure 3.35(b)), the SRCD signature changes and depending on the orientation the protein is dried in, the CD spectrum is shaped accordingly. However, in the half-dried samples, the signals are quite large, which means that could reflect LD instead of CD results, which is an artefact that needs to be

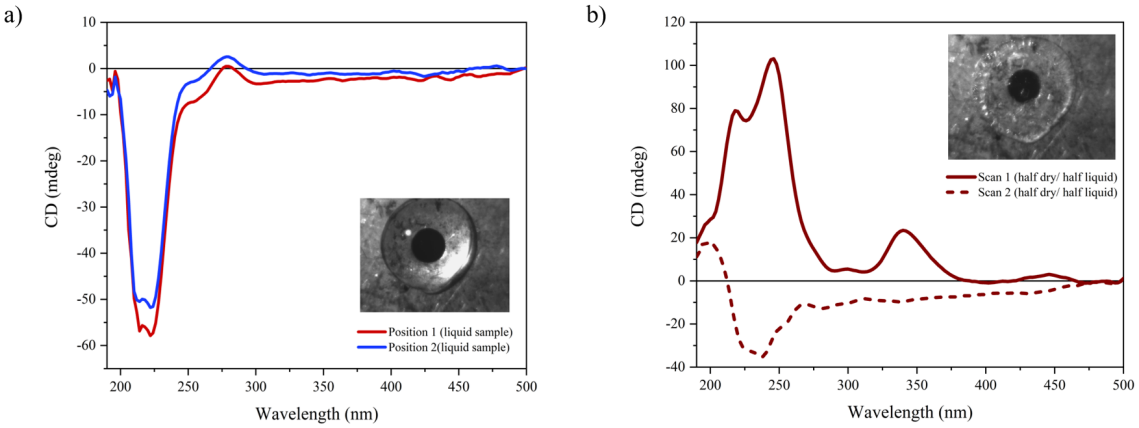

Figure 3.35: SCRD spectra of DX-array$_{(3')}$+cytb_{562} as the sample goes from **a)** liquid form to **b)** half dry/ half liquid state.

considered since it is directly associated with drying. Therefore, this fact needs to be taken into account when the films are prepared for oriented SRCD studies in order to ensure the proper conditions for the experiment that will lead to better quality of results.

3.3 Discussion and conclusion

In this Chapter, different protein variants of cytochrome b_{562} were introduced which results in diverse orientations upon hybridisation to the DNA nanostructure. To achieve this, a new DNA nanostructure was introduced, the DX-array, since a number of challenges were faced in Chapter 2 with the 2D DNA nanogrid. The formation of the DX-array proved to be successful, as well as its hybridisation with cytb_{562}, according to agarose gel results and AFM analysis.

The complex system (DX-A or array with protein attached) along with its reference samples (wild-type protein, azF ssDNA modified protein, DX-A tile and array) were tested both in liquid using conventional CD, and in films using SRCD spectroscopy and the CDi method. Various challenges were faced throughout these experiments. For the experiments in solution, the greatest complication was the buffer interference at the far UV and, as a result, important information about protein's characteristic peaks is missing. Based on the spectra after the subtraction of the DX-A and the DX-array, leaving just the CD of the protein-ssDNA, is difficult to draw any conclusion about any conformational change of the protein upon its hybridisation to the tile or array. This is mainly due to the fact that there are not enough information if the folding of DNA has been affected in the far-UV region, where the protein's and DNA's chromophores absorb. The DNA

retains its positive band at 280 nm, which comes mainly for the DNA, after protein attachment and it's not present after the subtraction and, it could be said that the DNA folding does not change. However, the cut-off at 210 nm does not provide further information about the DNA and therefore, it does not allow a confident protein comparison. The different variants of cytb_{562} also showed diverse behaviour when they were modified with either just an ssDNA or a bigger DNA nanostructure, leading to the conclusion that their orientation on the DNA can be distinguished using CD experiments in solution. On the other hand, for the experiments with the films, the biggest challenge was their preparation. Homogeneous films with the proper thickness and transparency are preferred in order to avoid any artefacts that arise due to sample inhomogeneity. We successfully determined the homogeneity of prepared DX-array films through CD mapping, and observed the first oriented SRCD spectra of cytb_{562} immobilised to the DX-A tile or array. The protein signatures compare quite well to the native protein in solution. However, the intensities of the bands were scalable for different positions of the same film, indicating thicker or thinner regions with greater or lower number of proteins present, respectively. Visual inspection of the films showed regions with aggregation or crystallisation of the sample which could result to artefacts which could also derive from the dehydration effect. Using the CDi method, these areas can be excluded from the sample and this is an advantage of this method compared to other oriented SRCD techniques that average the signal over a larger area.

Even though the CDi method and the film preparation techniques have been established, there are still several issues that needs to be overcome. The arrays with the protein are not always stably absorbed on the fused silica crucible, and removing the excess salt and protein through washing is not practical and reproducible. Multi-layers are also likely to form which lead to averaged readouts rather than single orientation information. The biggest challenge is the tendency of the film to create the coffee ring effect, where the sample is concentrated mostly at the edges of the film. In order to investigate whether these issues could be overcome, a new modified system is introduced and discussed further in the following Chapter.

Chapter 4

DX-array immobilisation on fused silica surface using fluorous effect

In the previous Chapters, various challenges were faced in the attempt to create suitable films with the right thickness and homogeneity, factors that are pivotal for the oriented SRCD studies in films. The coffee ring effect and the partially crystallisation of the protein samples at the edges of the film were some of the biggest challenges. The CD*i* method is capable of overcoming these problems since specific areas of the film can be probed by excluding regions where anomalies are present. Nevertheless, especially for the protein films, it was noticed that most of the sample and, by extension, most of the information was clustered at the edges, and upon dehydration the crystallisation of protein was inevitable. Therefore, the discovery and introduction of a new immobilisation approach, which would allow immobilising the sample of interest on the fused silica crucible in a more controllable way, was essential.

A potential approach would be the modification of the fused silica surface with microarrays, which allows biomolecular interactions using small amounts of samples. This method involves the sample to be covalently bounded to the substrate and hence, the molecule of interest needs to be chemically modified, which is not easy in most cases.[135,136] A recent, simpler method to attach biomolecules to solid supports came from Flynn *et al.*. Based on the work of Nichola Pohl, which demonstrated the formation of microarrays relying on the non-covalent fluorous effect, Flynn *et al.* made a step forward by attaching fluorous-tagged DNA onto fluorinated-solid supports.[137,138] The advantages of this approach are (a) the high resistance to non-specific binding, (b) strong and specific affinity of the fluorinated substrate to biomolecules including fluorous modifications and most importantly, (c) the complete removal of surface bound biomolecules using simple washes with

solvent (methanol) and the complete regeneration of the surface, allowing for the immobilisation procedure to be repeated.[139]

Herein, the fluorous effect approach will be exploited by modifying the fused silica crucible with perfluoro-carbon chains, as well as the DX-array. This modified approach has a three-fold advantage: (a) the grid will be stably attached and in a confined area, allowing the surface to be washed thoroughly to remove excess salt, DNA and protein; (b) the grid will be absorbed in a much more controllable way where the surface exposing the fluorous tag will be oriented towards the fused silica surface selectively; (c) the DX-array can be manipulated on the surface and is only released under stringent conditions.[138] This new arrangement twill allow to orient the conjugate onto the surface more precisely, to obtain cleaner systems, and, as a result, obtain SRCD signatures of superior quality. Therefore, the aims of this Chapter are the following:

1. Modification of the DX-array and the fused silica crucible with perfluoro-carbon chains.
2. Evaluation of successful immobilisation of the fluorous modified DNA nanostructure to the fluorinated fused silica crucibles.
3. Execution of SRCD experiments in order to obtain the preliminary data of the system before and after protein hybridisation.

4.1 Fluorous Chemistry

Fluorous chemistry involves the studies of the structure, composition, properties and reactions of highly fluorinated molecules. The term “fluorous” refers to a class of perfluorinated organic compounds in which all of the hydrocarbon backbones are substituted with fluorine atoms, leading to the general formula C_nF_{2n+1}.[140] Horváth and Rabai first coined this term as opposed to term “aqueous”.[141] A few years after the first use of term ”fluorous”, Gladysz and Curran suggested a formal definition as ”of, relating to, or having the characteristics of highly fluorinated saturated organic materials, molecules, or molecular fragments”.[142]

Highly fluorinated or perfluorinated compounds have a tendency to exclude themselves from both aqueous and organic phases. The answer behind this innovative feature of fluorous chemistry and the fluorous interactions is the unique atomic properties of fluorine atom exerted when introduced to organic molecular frameworks. The unique characteristics derive from fluorine’s high electronegativity, steric profile, and low polarisability.[143] When fluorine is bond to carbon atom, it

forms the strongest bound to a in organic chemistry, and the stability of C-F bond is increased as the number of fluorines bound to the same carbon increases.[144] The secret to this stabilisation is the overlap between fluorine 2s and 2p orbitals with the corresponding carbon orbitals. The C-F bond is polarised from the sp^3 carbon (δ^+) to the fluorine (δ^-), ih the opposite direction from C-H bond, due to the high electronegativity of fluorine. Despite the higher electronegativity though, C-F bond involves lower van der Waals interactions due to low polarisability.[144] According to Pauling's principal, the fluorous-fluorous non-covalent interactions are also stabilised and can give significant energy distribution to the structures due to the low polarisability of fluorine.[145]

In contrast to hydrocarbon chains, fluorous chains are both hydrophobic and lipophobic, avoiding unfavoured interactions of fluorine atoms with other elements, the so-called "fluorous effect".[142,146] Exploiting the selectivity of fluorous-fluorous interactions, Curran *et al.* introduced the fluorous synthesis, which included tagging a molecule with highly fluorinated tags which was then used as analogs of solid phase synthesis and later on for fluorous solid phase extraction and fluorous chromatography.[147,148] Since then, a vigorous research has relied on fluorous tags for many applications in synthesis and separation technologies, in catalysis and in novel technologies, such as microarray developments.[137,149,150]

An organic compound can be modified to a fluorous analog by incorporating fluorous tags which include a perfluorocarbon group (R_F), a suitable spacer, such as a hydrocarbon spacer, and the organic compound of interest whose reactivity remains dominant. For this kind of tagged molecules, the new incorporated fluorous region is usually referred to as a *fluorous ponytail*.[151] This novel approach of using perfluorocarbons requires just a small single perfluorocarbon tail, most commonly C_8F_{17}, to be attached as a tag to a compound. This method was demonstrated in such a way to ensure stable immobilisation from water solutions into a perfluorocarbon-coated solid surface. Further to this and unlike traditional microarrays, the perfluorinated substrates could be rinsed and re-used multiple times.[139] That makes the fluorous effect a very strong candidate for an effective alternative to the covalent immobilisation approach.

4.2 Results and discussion

4.2.1 Fabrication of fused silica surface using fluorous tags

As mentioned earlier in this Chapter, many challenges were faced in the previous two Chapters especially during the formation of the films. The DNA nanostructure

is not stably absorbed on the fused silica surface, and removing excess salt and protein through washing is not feasible. Furthermore, multi-layer systems are likely to form and in general, the system is prone to the "coffee ring" effect and the partially crystallisation of the protein upon dehydration. Therefore, it was essential to devise an alternative approach of depositing the sample to the solid surface, which would also require small sample volumes. A potential solution to these issues, which is explored in this Chapter, is the immobilisation of the DNA nanostructure to the solid surface using fluorous tags. This method would provide a stable system that can be washed and most importantly, the surface can be re-used multiple times.

For performing the SRCD experiments relying on the fluorous effect approach, it was essential to modify the fused silica crucible which is used in all the experiments presented in this Thesis. Therefore, a fused silica crucible (15 mm of internal diameter, Linkam) was sent to our collaborator, Alasdair Douglas from the University of Glasgow, in order to proceed with the fluorous modification. The quartz crucible was first cleaned thoroughly to remove any organic contaminants and particles by sonication in acetone, followed by isopropanol (IPA) rinses and oxygen plasma treatment. In order to create specific patterned areas, photolithography was carried out and finally the substrate was silanised with the Heptafluoro-1, 1, 2, 2-tetrahydrodecyl trimethoxysilane (FTDS).

Figure 4.1 presents the fluorinated patterns that were selected to be created on the crucible. The first fluorous array (left) consists of multiple 50x50 μm^2 mini squares as a sensing area extended along the surface of the substrate, whereas the second one (right) has a single larger square of 3x3 mm^2.

Usually, the silanisation of the surface is investigated by determining the hydrophobicity of the FTDS-modidied substrate using static contact angle measurements. In the case of the fused silica crucible here, our collaborators characterised the FTDS silanised crucibles with mini squares pattern presented in Figure 4.1. The characterisation showed that the contact angle measurement was in accordance with all the reported values for FTDS modified surfaces, demonstrating typical contact angle measurements of 114°.[142,152] On the other hand, the square becomes distinguishable for the second crucible (single square) by depositing a few drops of water, since no water covers the hydrophobic square. Therefore, it was concluded that the silane was deposited onto the surface.

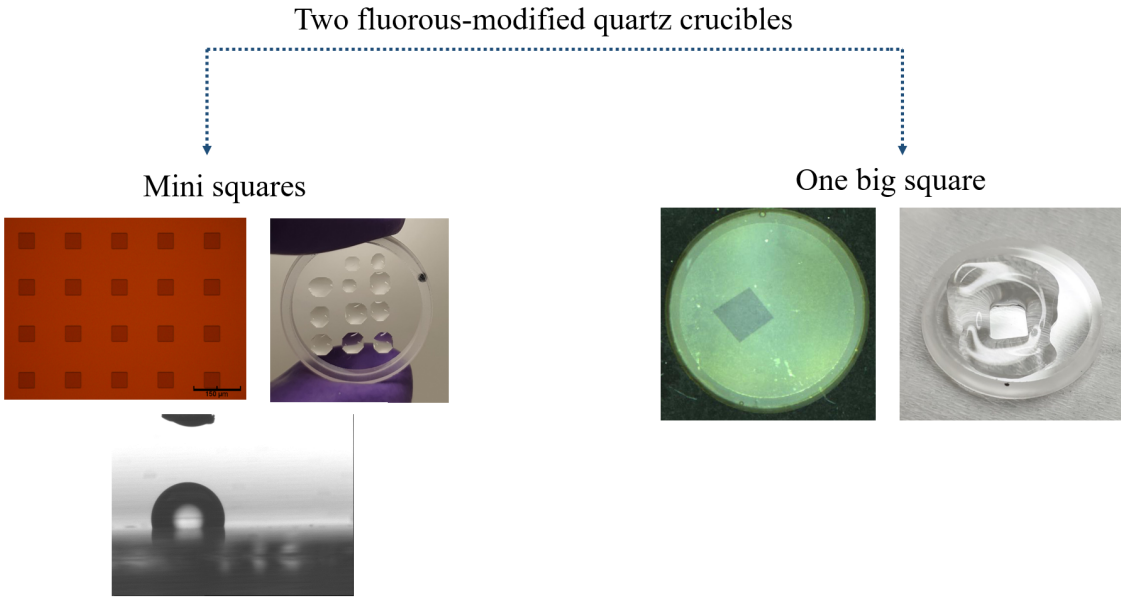

Figure 4.1: Pictures show the FTDS- modified surfaces on the quartz crucible. Multiple mini squares of 50x50 μm^2 or big square of 3x3 mm^2. The positioning of a water drop on the hydrophobic surface of the fused silica with the multiple mini squares is also present.

4.2.2 Formation of fluorous-tagged DX-array

Two different DNA nanostructures have been used during this project, the 2D DNA nanogrid and the DX-array. However, only the latter forms very well in solution. Therefore the DX-array was selected to get modified with the fluorous tags for its immobilisation to the fluorinated quartz crucible via fluorous interactions.

As mentioned in Section 3.1, Rusling's original DX-array was modified in such way to create an anchoring point for protein hybridisation. This was accomplished by modifying one of the tiles used for the generation of the DX-array. In this case, the middle strand of the DX-A tile was modified (Figure 3.3). For the fluorous tag modifications though, the DX-B tile was chosen in order to avoid making the system more complicated.

First, the most appropriate location for the modification needed to be addressed so that the fluorous tags would be facing outwards, away from the extended lattice and in opposite direction of the protruding anchoring point of the DX-A tile. In other words, ideally the fluorous tag would be facing down towards the crucible while protein's hybridisation point will be facing up, away from the crucible and the DX-array surface. Therefore, the 5' terminal of the DX-B middle strand, where the fluorous tag would be introduced, was moved by 21 bases. The original 5' terminal position of the DX-B tile, as well as the modified DX-B tile are presented in Figure 4.2 (a) and (b), respectively,. The strand that has been modified is

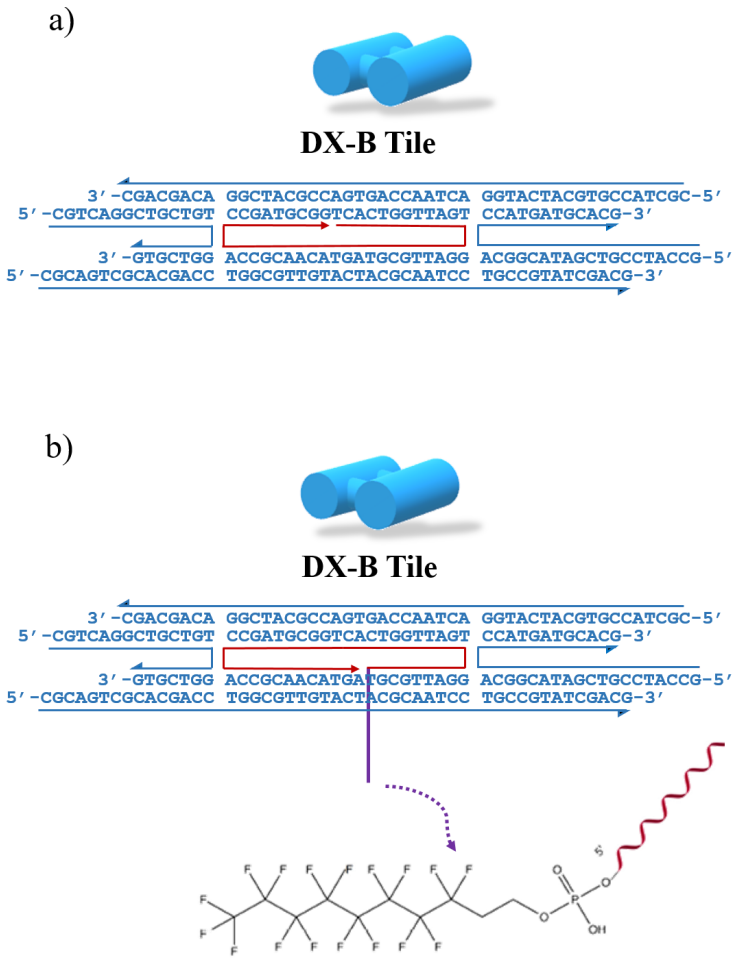

Figure 4.2: Schematic drawing including the sequences of DX-B tile **a)** Original design. The red strand indicates the strand that will be modified. **b)** DX-B tile after alteration of the middle strand's 5' end (red line) and after modification with the fluorous chain (purple solid line). An example of fluorous chain is pointed out with a purple dashed arrow.

indicated with red colour. After this alteration, the middle was further modified by introducing the fluorous tag at its 5' end (purple solid line in DX-B tile, Figure 4.2(b)).

The most common fluorous-phase tags used for the immobilisation of biomolecules onto fluorinated solid surfaces are the mono-C_8H_{17}. Nevertheless, Flynn reported that branched fluorous tags (bis-C_8H_{17}) can lead to stronger fluorous-fluorous interactions between the fluorinated biomolecule and the solid surface.[153] It was considered that the branched-fluorous-tagged oligonucleotides showed an increased binding capability due to the increase in their distance from the surface. The further the molecule is from the surface, the better, since it is far from the hydrophobic surface and permeates greater flexibility to the molecule. In addition, another factor that plays an important role in the increase of binding activity is the presence (or absence) of a linker between the fluorous ponythail and the carbohydrates. In

the same research work, mono-C_8H_{17} fluorous tagged oligonucleotides with and without linker (Hexaethyleneglycerol , HEG, linker was used) were compared and it was found that the fluorescence activity value for the oligonucleotide containing the linker was significantly higher than the other. Similar to the branched-fluorous-tagged molecules, it is likely that the oligonucleotide without the linker is too close to the hydrophobic fluorinated surface and therefore, it is sterically prevented from participating in the hybridisation reaction. As such, the presence of a HEG linker is beneficial for the fluorous- tagged molecule since it increases its flexibility.[153]

Taking into account all the observations regarding the effect of the amount of the fluorous content and the presence (or absence) of a linker on the immobilisation performance of the fluorinated molecule to the fluorinated surfaces, three different modifications were selected in order to be used for the modification of the tile's middle strand. These modifications are: (i) mono-C_8H_{17} with four thymidines as a linker instead of HEG, (ii) mono-C_8H_{17} with a HEG linker and (iii) bis-C_8H_{17} with four thymidines as a linker instead of HEG. Furthermore, the DX-array is a very large molecule, therefore, in order to ensure its successful immobilisation to the fluorinated crucible, two more strands will be modified on their 5' end with mono-C_8H_{17}-fluorous tags with four thymidines as a linker. These two strands are next to the middle strand (DX-B3), the side DNA strands DX-B1 and BX-B2. In Figure 4.3, the DNA strands and their 5'-end fluorous modifications and respective linkers are presented. Dr Andrea T. Sender from Glenn Burley's group (University of Strathclyde) carried out the synthesis and HPLC purification of the fluorous-tagged oligonucleotides.

Different combinations of the fluorous modified strands were tested for the formation of the individual DX-B tiles as well as the DX-arrays. However, it was not possible to investigate all combinations, due to time limitations. The most probable combinations to increase the stability of the DNA nanostructures on the fluorinated surfaces were chosen. The DX-B tile and DX-array were formed in solution following Protocol F2 described in Section 6.10. The formation of the DNA nanostructures was then monitored using 1.0% agarose gel electrophoresis.

Observing the gel presented in Figure 4.4, it is clear that the DX-B tiles (4 μM) are successfully formed, since each respective band is visible on a higher level than the control sample (1 μM), which consists of just two strands, DX-B3 and DX-B4, creating a duplex. However, there is no clear difference among the DX-B tiles with different fluorous content on their structure. The DX-array (2 μM) also seems to form as expected, since there is a band at the top of the gel indicating the formation of a bigger DNA structure than the individual tiles, which is in accordance with all the agarose gels shown in this thesis. Each of the followings gels was run at different dates. The names used for each DX-B tile and DX-array sample on the

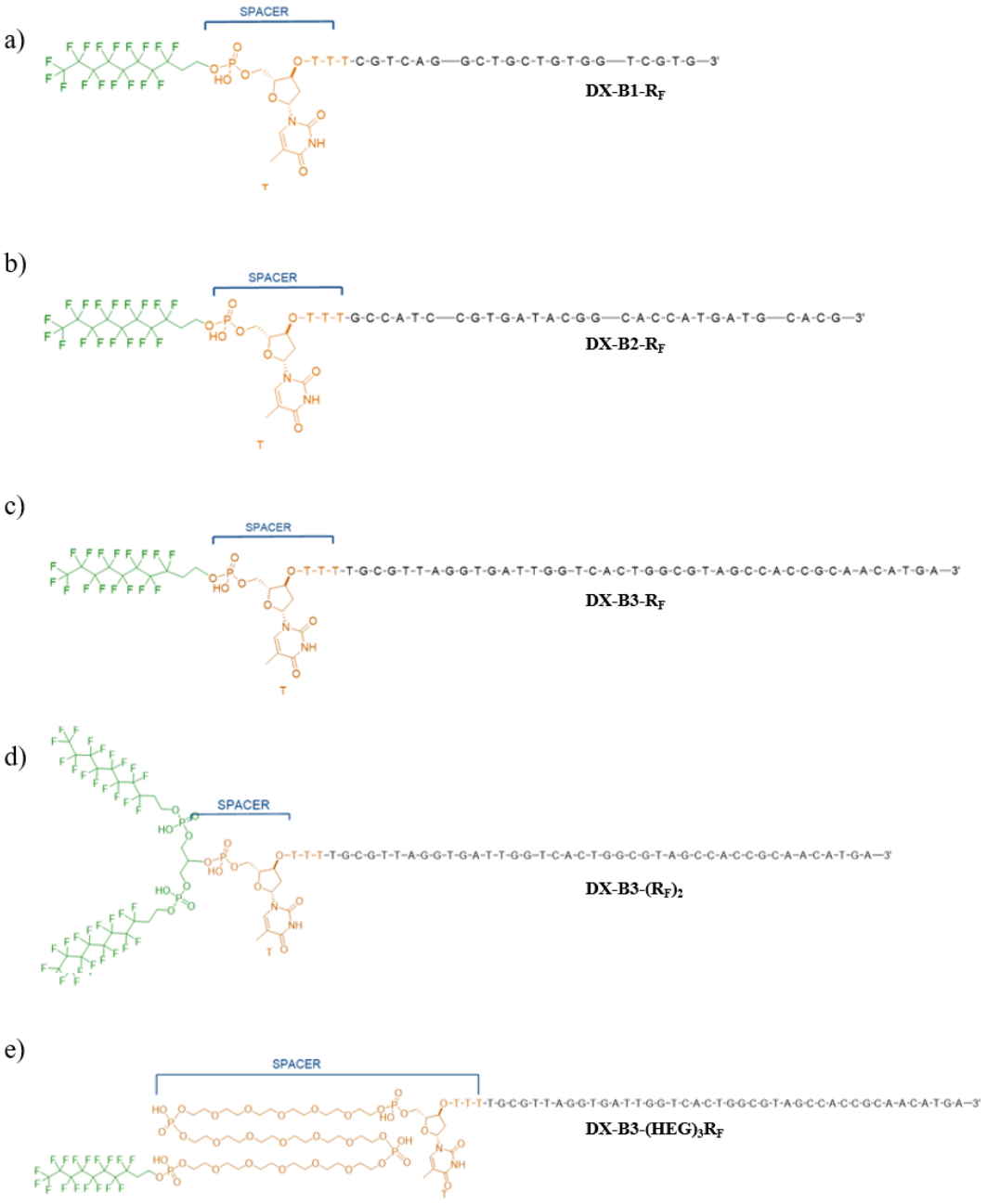

Figure 4.3: Schematic of the different branch fluorous-phase tags attached to the 5'-end of the DX-B tile DNA strands. **a)** DX-B1 strand with mono-C_8H_{17} and 4 Ts as a linker. **b)** DX-B2 strand with mono-C_8H_{17} 4 Ts as a linker. **c)** DX-B3 strand with mono-C_8H_{17} 4 Ts as a linker. **d)** DX-B3 strand with bis-C_8H_{17} 4 Ts as a linker. **e)** DX-B3 strand with bis-C_8H_{17} 4 Ts as a linker.

gel description are shown in Table 4.1, along with the fluorous modified strands that were used for their formation.

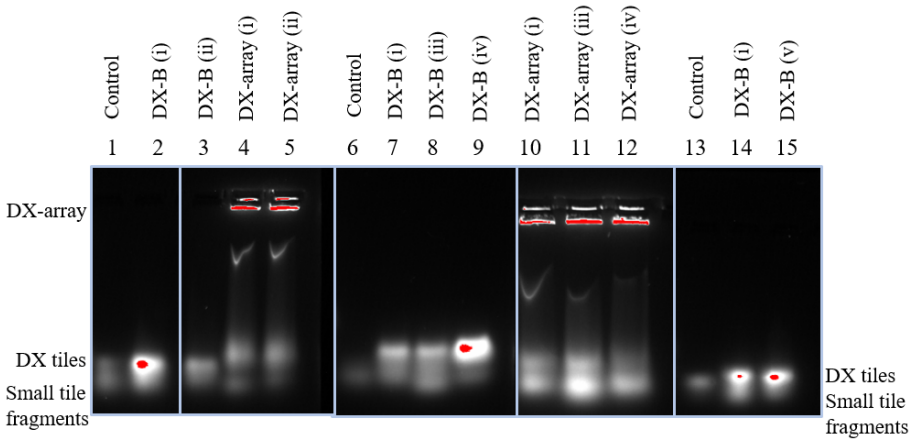

Figure 4.4: 1.0% agarose gel of the fluorinated DX-B tiles and DX-arrays. The differences among the samples can be found in Table 4.1.

Fluorous modified DX-B tile	Fluorous modified DX-array tile	Fluorinated DX- B strands
DX-B (i)	DX-array (i)	DX-B1-Rf DX-B2-Rf DX-B3-$(HEG)_3$-Rf
DX-B (ii)	DX-array (ii)	DX-B1-Rf DX-B2-Rf DX-B3-$(Rf)_2$
DX-B (iii)	DX-array (iii)	DX-B1-Rf DX-B2-Rf DX-B3-Rf
DX-B (iv)	DX-array (iv)	DX-B1-Rf DX-B2-Rf
DX-B (v)	-	DX-B3-$(Rf)_2$

Table 4.1: DX-B tiles and arrays run on the 1.0% agarose gels with the respective fluorous modified strands which are included in each of these samples.

4.2.3 Immobilisation of the fluorous-tagged DX-array on the fluorous modified fused silica surface

The final step prior to the SRCD experiments was the immobilisation of the formed fluorinated DX-array on the fluorous modified fused silica crucible. The steps for the immobilisation procedure can been seen in Figure 4.5.

Briefly, the quartz crucibles were fabricated using standard photolithography and the pattern created was either fluorous modified mini squares or a single big square (Figure 4.1) surrounded by a non-fluorous background. Following, the fluorous-tagged DX-array that was formed in solution, with or without protein attachment, was incubated onto the crucible for 30 minutes minimum in a humidity chamber

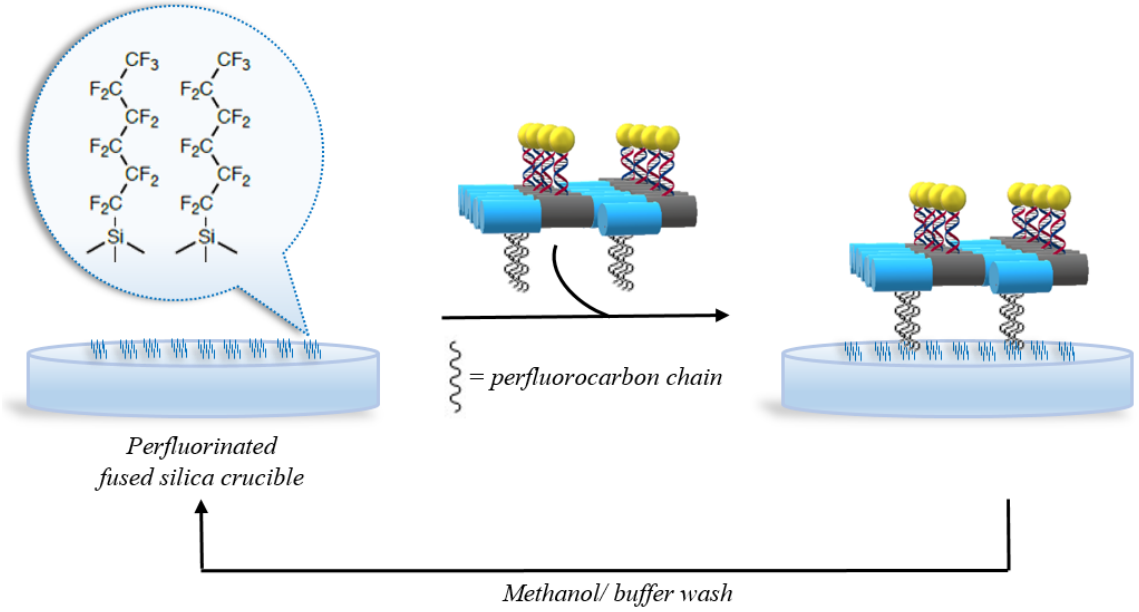

Figure 4.5: Schematic shows the immobilisation of the DX-array onto the photolithographically defined fluorous arrays of the fused silica crucibles. DNA strands tagged with fluoro ponytails will strengthen the absorption of the DX-array on the surface. After a simple wash with 50:50 methanol/ buffer, the perfluorinated fused silica crucible is regenerated and ready for re-patterning.

in order to prevent any evaporation. A washing step was added, to remove non-specific bound proteins, staple strands or salt. Finally, the crucible was dried very carefully and slowly under a stream of nitrogen. As mentioned earlier, the immobilisation is fully reversible, after 50% methanol in hybridisation buffer wash, enabling directed surface patterning, regeneration and re-patterning of surfaces without any associated degradation of immobilisation efficiency or disruption of Watson-Crick base-pairing.

In order to detect the presence of the DX-array on the quartz crucible, fluorescence microscopy was carried out and GFP protein was hybridised on the DX-array (1.5:1 protein/ DNA ratio). The DX-array with the fluorous modifications DX-B1-R_F, DX-B2-R_F and DX-B3-$(HEG)_3$-R_F was used, and the array formation and protein hybridisation were carried out in solution. Many experiments led to misgivings about the presence of the DX-array on the surface crucible and, therefore, the assessment of this argument was very important, as discussed in Section 4.2.4 below. Access to Diamond Light Source Beamline B22 was granted in order to examine samples of DX-array attached with GFP on both crucibles using fluorescence microscopy (ZEISS Axio Imager 2, Carl Zeiss Microscopy). Unfortunately, the amount of the available sample was not adequate, therefore it was tested only in the crucible with the mini square patterns. If the immobilisation of the DNA through fluorous interactions worked in this crucible, then this result would apply to the other crucible as well. An image of the DX-array with GFP attached as

captured using the fluorescent microscope is presented in Figure 4.6. It is clear that the fluorous modified DX-array can be immobilised on the fluorinated squares, since the protein is visible and also the proteins are in line, according to the design of the complex. According to AFM, the repeat spacing of the proteins is 32 nm. In an attempt to try and measure the distance between two proteins (two dots), the distance was much bigger (1.5 μM). However, this might not be real, since there is not any other underlying structural features observed in earlier stages that would lead to this pattern, and there fore it can be concluded that the DNA is present on the crucible. GFP is sticky on fused silica surfaces and it needs to be washed thoroughly in order to be removed completely. In this case, the area between the squares is not masked and, hence non-specific binding cannot be avoided.

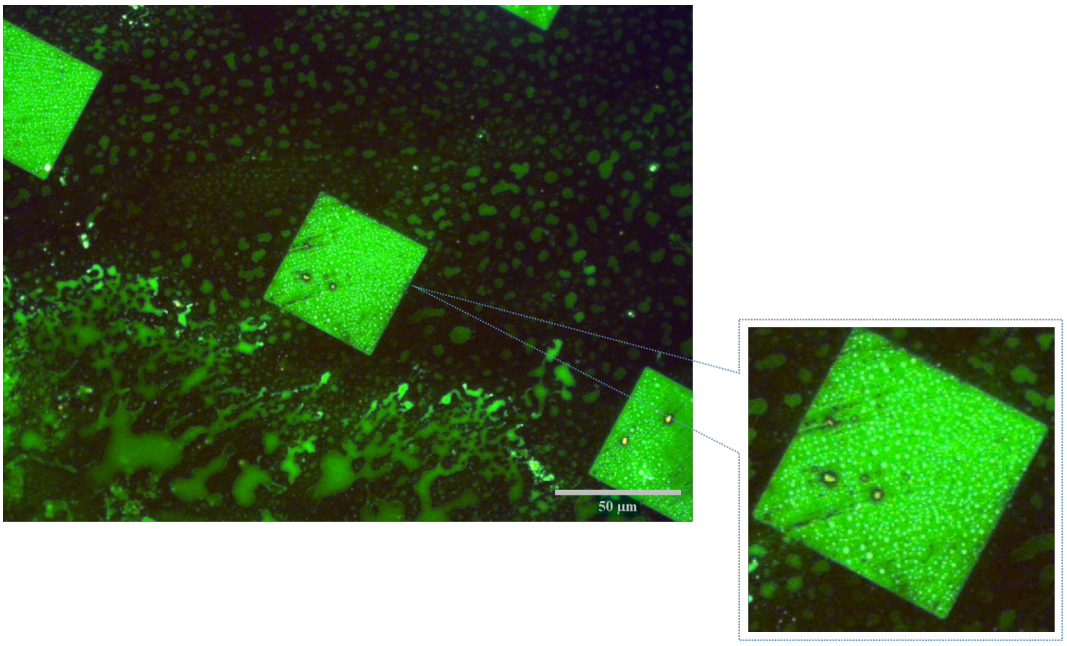

Figure 4.6: Image of the DX-array with GFP attached (1:1.5 ratio) on the quartz crucible after immobilisation fluorous effect as it was captured using fluorescence microscope (50x objective was used).

4.2.4 CD and SRCD measurements of the fluorous-tagged DX-array on fused silica surfaces

SRCD experiments of the DX-array were recorded at B23 beamline at Diamond Light Source during the allocated beamtime SM24438-1, which was the final one for this PhD studentship. CD*i* experiments were carried out after the immobilisation of the DX-array (with or without protein) onto the fused silica crucible using fluorous effect, an alternative sample preparation method, which has the potential to overcome all the challenges faced in the previous Chapters. Therefore, the aim of this Section is to investigate whether this new approach is suitable for

discriminating DNA or proteins which are hybridised in DNA nanostructures by SRCD or CD*i*.

As mentioned in Section 4.2.2, various DX-arrays were formed which included different number of fluorous ponytails in order to assess whether there was a significant stability difference among the systems on the perfluorinated crucible. Although all the DX-arrays were intended to be tested, only the DX-array-$(HEG)_3$-R_F was carried out using both crucibles templates due to time constraints. The fluorinated DX-array-$(HEG)_3$-R_F was chosen since it consists of 3 strands modified with the fluorous ponytails and it is assumed to be more stable on the fused silica, as shown by the fluorescence microscopy in Section 4.2.3. Therefore, in the SRCD experiments that follow, the standalone DX-array-$(HEG)_3$-R_f, which will be referred to as *DX-array*, as well as modified with the Q71 cytb_{562} (short axis, monomer) or Q50 cytb_{562} (long axis, dimer) variants attached were examined using the CD*i* technique.

DX-array

Initially, the crucible with the mini squares (50x50 μm^2) was used for the SRCD experiments of the DX-array and two different CD*i* experiments were run. 4x4 with 60 μm step size (covering an area of 240 μm x 240 μm) and 14x14 with 40 μm step size (560 μm x 560 μm) grid arrays were scanned and CD, as well as UV absorbance data were obtained and plotted as shown in Figure 4.7. The average of all CD and absorbance scans is shown in black solid line.

Both grid scans yielded in similar CD signal as expected, since the only difference was the size of the scanned area. The CD signature, and especially for the CD spectra presented in Figure 4.7 (b), seems to resemble the signature of the DX-array which was recorded in previous sessions in films and showed a negative peak at ∼220 nm and a positive at ∼275 nm. However, in this case, the positive peak is slightly shifted at 260 nm. But since the CD spectrum of a DNA is sequence and composition-dependent, maybe this small modification to the DX-B tile led to this shift. Moreover, the intensity of the CD signal is also weaker (∼3 mdeg) than the previous results (8 mdeg) when the sample was immobilised on the crucible as a film. This is also noticeable in the UV absorbance data in which the absorbance is below 0.1. Ideally, the absorbance should be greater than 0.5 in order to make sure that we record above the noise. However, it is not crucial to get a very strong signal for the DNA nanostructure since the aim is to record the CD of the protein when it is attached to the DNA.

The 2D maps plotted using the SRCD data from the 14x14 grid array scan at 220 and 258 nm, respectively, could give some information about the homogeneity of the sample deposited on the crucible. Both maps yielded very similar patterns,

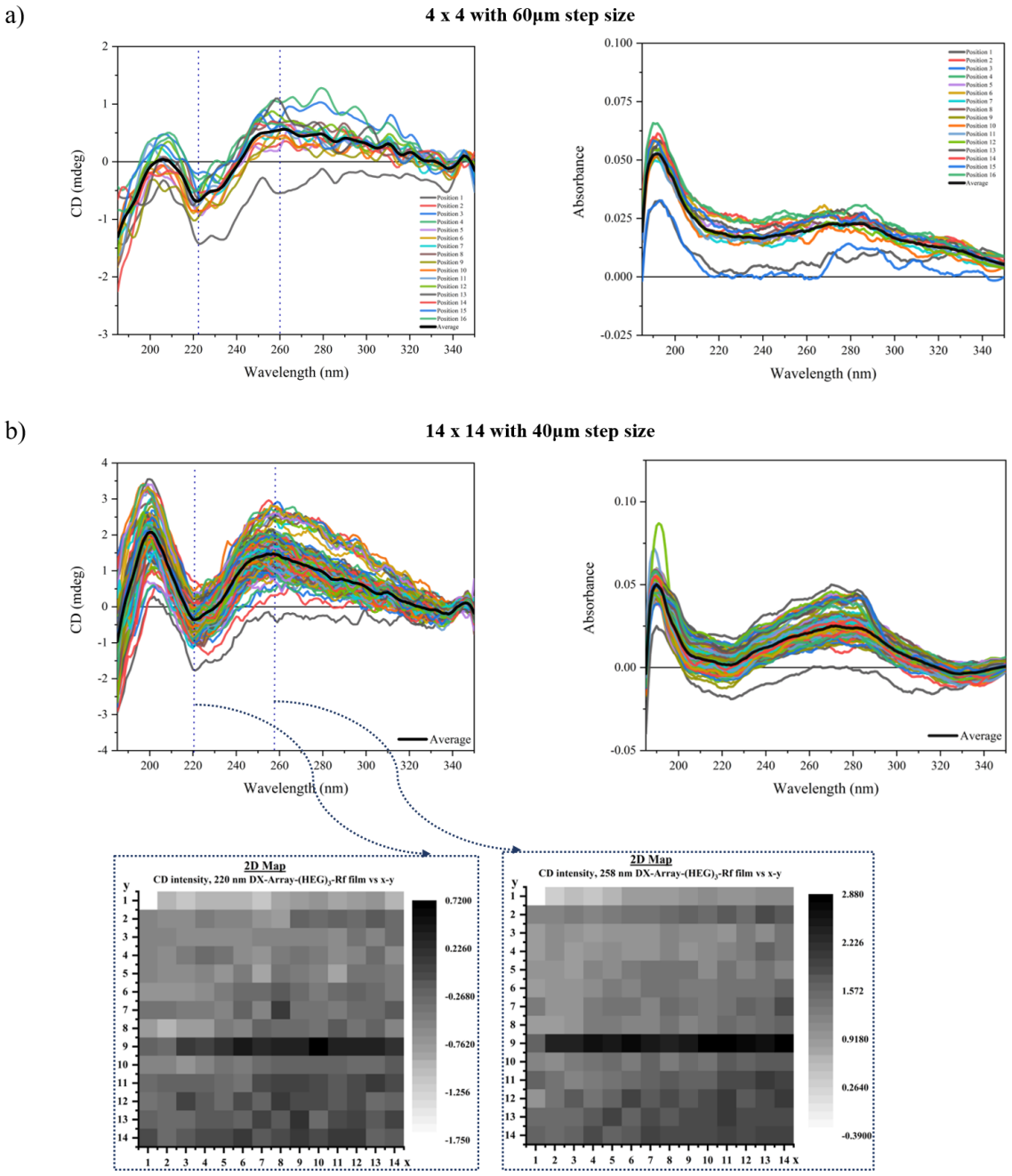

Figure 4.7: a) CD and UV absorbance data of DX-array as they were recorded during CD*i* experiment by scanning 4x4 grid array at intervals of 60 μm. **b)** CD and UV absorbance data of the same sample obtained by scanning 14x14 grid array at intervals of 40 μm. 2D maps of SRCD intensity at 220 and 258 nm, respectively, vs x-y; each square represents the spot mapped (40 μm^2) during the 14x14 grid array scan.

which show an adequately homogeneous sample. Nevertheless, it is surprising that there is a CD signal in almost every spot mapped, since not all of them should have DX-array sample. According to the pattern of the modified disk, DNA should be present only in the area of the mini squares where the fluorous interactions will take place upon the sample addition. Moreover, it is not expected that the beam will hit each mini square after every 40 μm step move. Therefore, this probably is a

consequence of the fact that the area between the fluorous modified squares is not masked in order to avoid any non-specific binding, and this outcome is explained. Another explanation though that could be possible is that the CD*i* mapped the CD features of the fused silica since there is no film created following this sample preparation approach, but only one molecule stacked to the surface and therefore it is is too dilute too get detected.

The 2D map of the single wavelength CD*i* experiments for the DX-array at 265 nm is presented in Figure 4.8, where a 20x20 grid array was scanned with 40 μm step size, corresponding to an area of 800 x 800 μm. This map could provide some useful information about the pattern of the sample on the quartz and, in general, information about the homogeneity at the surface of the crucible. However, a specific pattern is not observed, and the result looks very similar to the previous two colour maps. The map was plotted automatically using the OLIS program while the experiment was running. This confirms that the resolution is too low to see the distinct pattern.

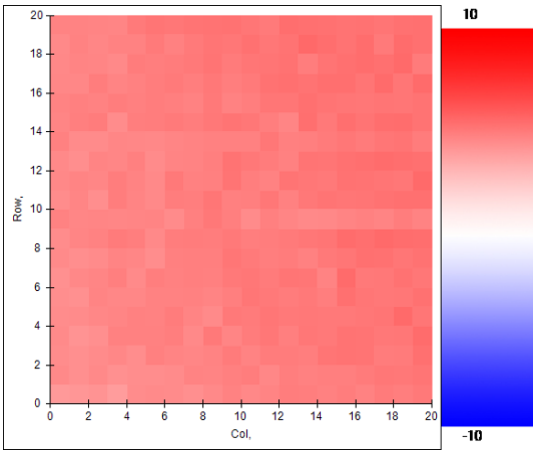

Figure 4.8: 2D CD map of single wavelength CD*i* experiment for the DX-array at 265 nm. A 20x20 grid array was scanned of 40 μm step size.

The DX-array was then tested in the crucible with the single big square (3x3 mm^2) and CD*i* experiments were performed for scanning a 4x4 grid array with 70 μm step size and a 10x10 grid array of 60 μm step size. In contrast to the previous crucible, by using the single square, it was expected that the sample would bind to the whole fluorous surface and the CD signal would be present in any spot mapped on the modified area. However, according to the data shown in Figure 4.9, the CD signatures do not feature the characteristic peaks of DNA and they are very different in comparison to the previous crucible because most likely one molecule on the surface is too dilute for detection. Therefore, it seems that since we are not creating a film with multiple layers with specific thickness and transparency, it is difficult to get a signal for the DNA, which is weak in general, as mentioned earlier.

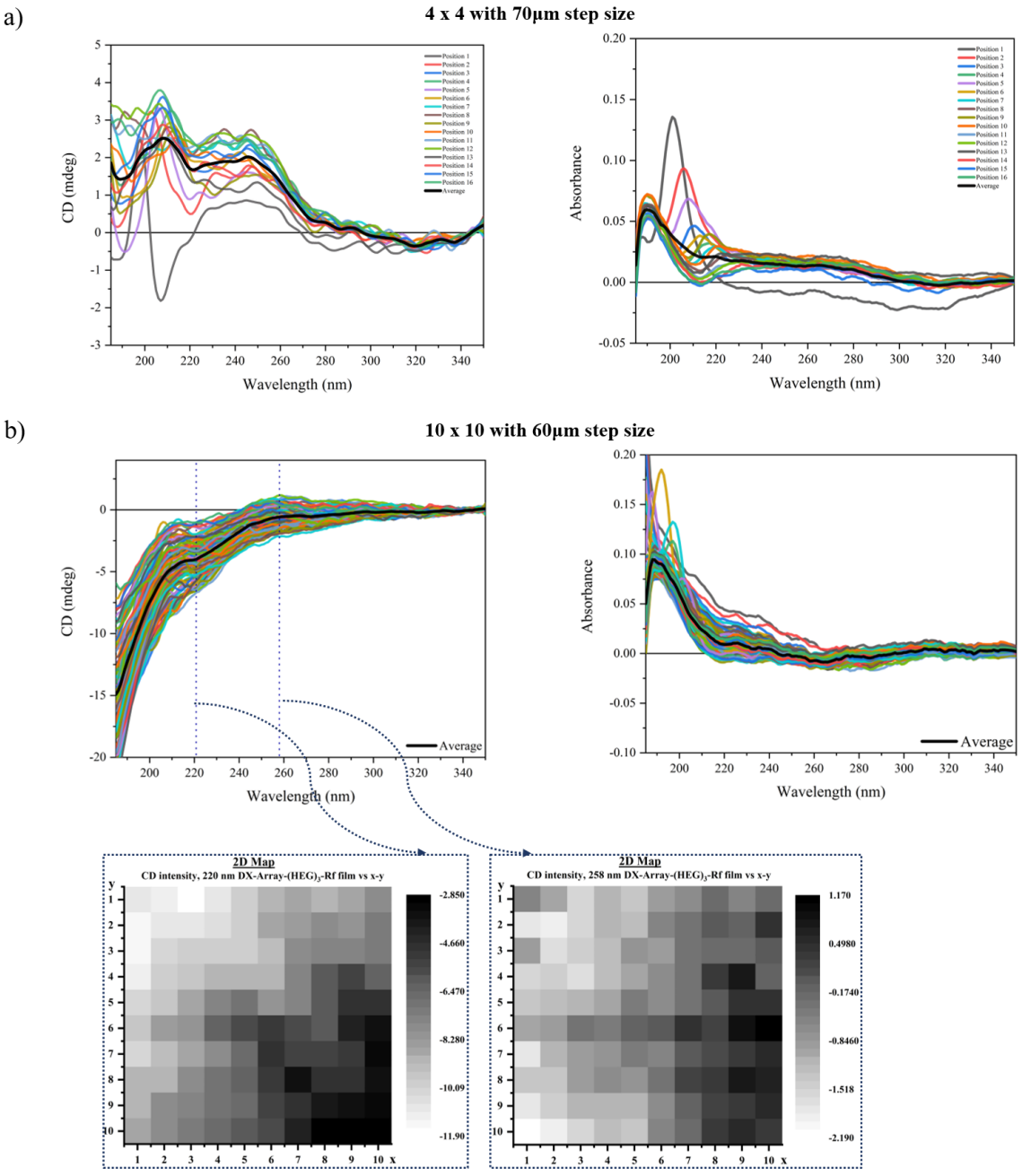

Figure 4.9: a) CD and UV absorbance data of DX-array as they were recorded during CD*i* experiment by scanning 4x4 grid array of 70 μm. **b)** CD and UV absorbance data of the same sample obtained by scanning 10x10 grid array of 60 μm. 2D maps of SRCD intensity at 220 and 258 nm, respectively, vs x-y; each square represents the spot mapped (60 μm^2) during the 10x10 grid array scan.

In terms of concentration and structure, homogeneity would be expected across the modified square. Nevertheless, it seems that in case the DNA is present but weak enough to complicate the recording of CD, the DNA has not bound to the whole surface, since there are areas with some signal and others with almost no signal, according to the 2D maps (Figure 4.9 (b)). This is also noticeable at the 2D map which was plotted after CD*i* at single wavelength (220 nm, Figure 4.10). This

suggests that covering a large area is not straight forward, and covering multiple smaller areas seems to be a better approach with self-assembled and extended DNA grids.

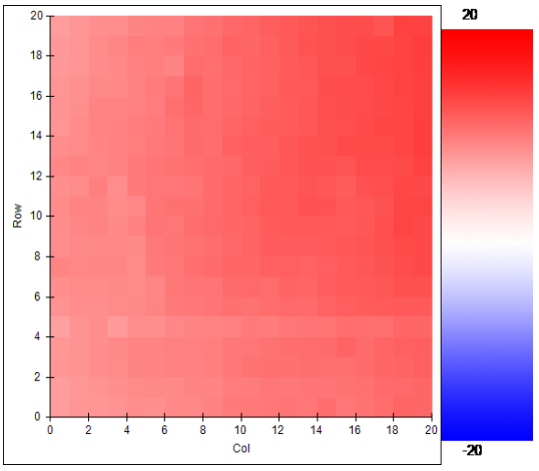

Figure 4.10: 2D CD map of single wavelength CDi experiment for the DX-array at 220 nm. A 20x20 grid array was scanned of 40 μm step size.

DX-array with Q71 cytb_{562} variant

The next sample that was tested included cytb_{562}, and more specifically the short axis variant Q71 (monomer). As stated in Chapter 3, cytb_{562} is an α-helical protein and it is also a heme protein. The heme gives an induced CD at ~420 nm. However, the following experiments were focused on the wavelength region 180 to 350 nm because, considering the previous SRCD experiments, the signal for the heme is not strong enough to be recorded, and therefore, the protein-DNA region was only investigated.

Initially, the Q71-ssDNA was pipetted on top of the DX-array that was tested previously in order to see whether the protein would be able to immobilise to the DNA nanostructure by attaching to its complementary strand. The DNA protein ratio was 1:1. A 4x4 grid array was mapped with 60 μm step size using the crucible with the mini squares first. In Figure 4.11 (a), the results originate. The minima of the α-helical protein are not visible and, in general, the signal remains very low. Since the signal resembles more the standalone DX-array, it leads to the conclusion that the signal for the DNA is dominant. Therefore, in Figure 4.11(b), the data shown are configured after the subtraction of the DX-array from the respective sample with Q71 added on top. It is evident that there are no significant changes, except for the fact that the CD intensity becomes even weaker without revealing CD signature characteristic for this type of protein, but more like a β-sheet, which has a minimum band between 215 and 219 nm and a larger positive band between 195 and 202 nm.

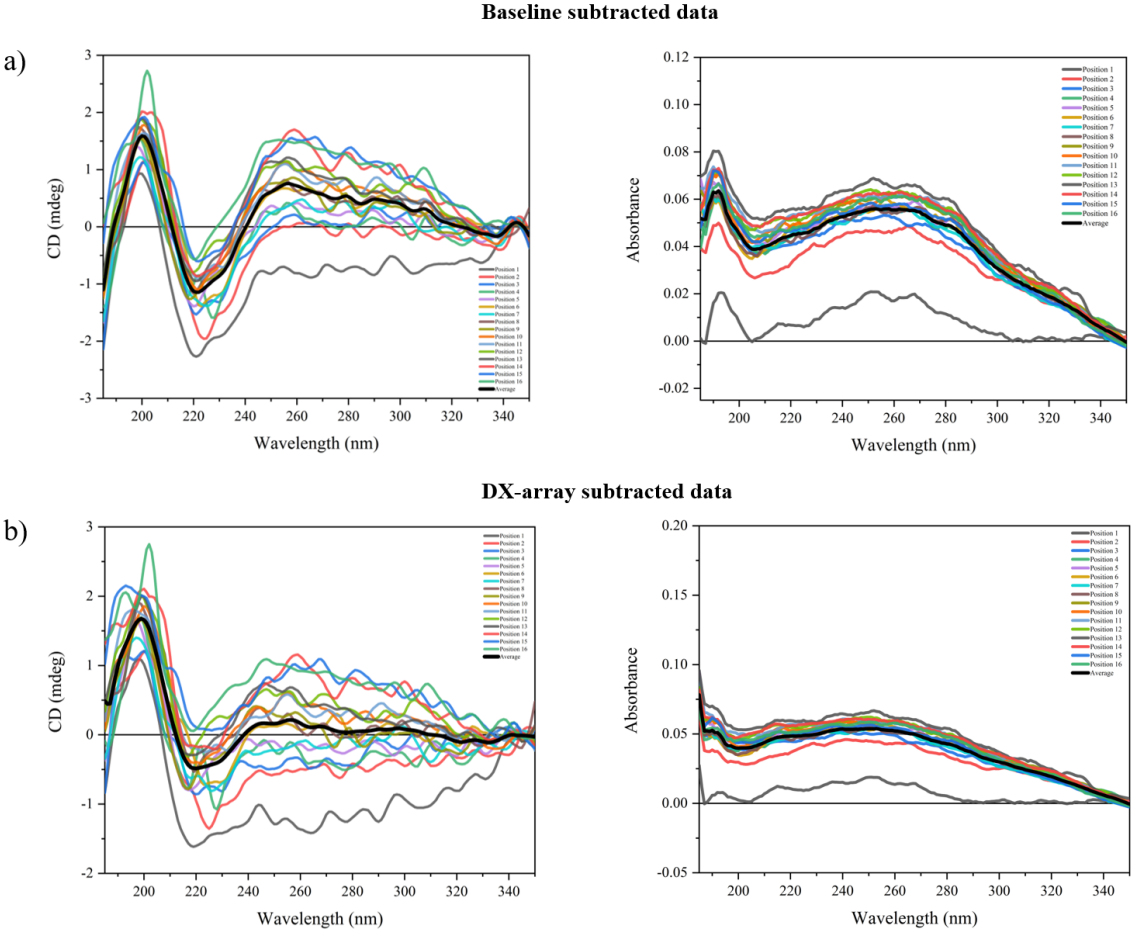

Figure 4.11: a) CD and UV absorbance data of DX-array as they were recorded during CD*i* experiment by scanning 4x4 grid array of 70 μm. **b)** CD and UV absorbance data generated by DX-array subtraction from the DX-array with Q71 attached sample.

It is important to note that the empty crucible (baseline), the disk with the DX-array or with the DX-array+protein were tested in different runs. This means that the crucible was removed from the stage and cleaned between runs so that the next sample will be deposited. However, it is extremely hard to ensure that the disk will be placed back to the stage at the exact same position, and as a result, it is uncertain if the different samples will be aligned. Therefore, when the subtractions are carried, this problem needs to be considered. The alignment, if in an ideal distribution, would need to be checked by LD, which could give an indication on the relative orientation.[154] This would then be used to orient the samples in the same way, measure CD and finally, use this to subtract the data. Nevertheless, it is not clear how feasible that would be in our case and due to time limitations, it was not possible to investigate it any further.

In addition, another sample of DX-array with Q71 cytb_{562} attached was prepared in solution (with protein/DNA ratio 1:1) and tested in the crucible with the mini squares in a 10x10 grid map with 60 μm step size. The baseline subtracted data

are presented in Figure 4.12 along with the CD maps, as plotted with the data collected at 208, 220 and 251 nm and the UV absorbance data.

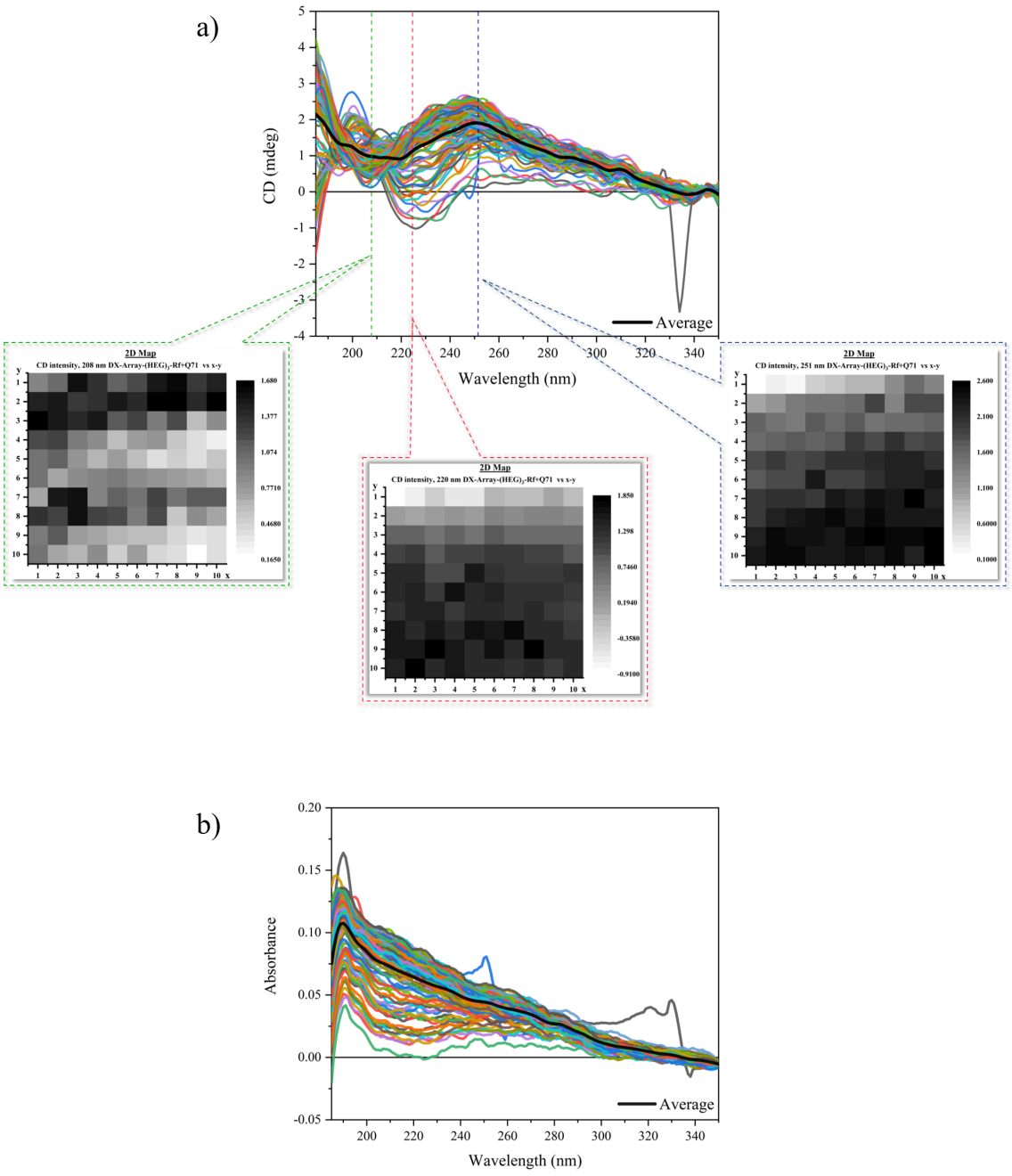

Figure 4.12: CD data of DX-array with Q71 attached as they were recorded during CD*i* experiment by scanning 10x10 grid array of 60 μm. 2D maps of SRCD intensity at 208, 220 and 251 nm, respectively, vs x-y; each square represents the spot mapped (60 μm^2) during the 10x10 grid array scan. **b)** Absorbance data of DX-array with Q71 attached

Here, the CD signal for the areas scanned show different behaviour compared to the earlier measurements. There are some negative bands at 220 nm but the majority of the scans are above zero at this wavelength. Therefore, the CD data are not clear enough for the protein region. However, the positive peak at 250-260

nm could indicate the presence of DNA for most of the scans but it is not clear enough whether the measurements were recorded above the noise or these spectra reflect the fused silica surface. The 2D colour maps also show a different pattern compared to the previous results.

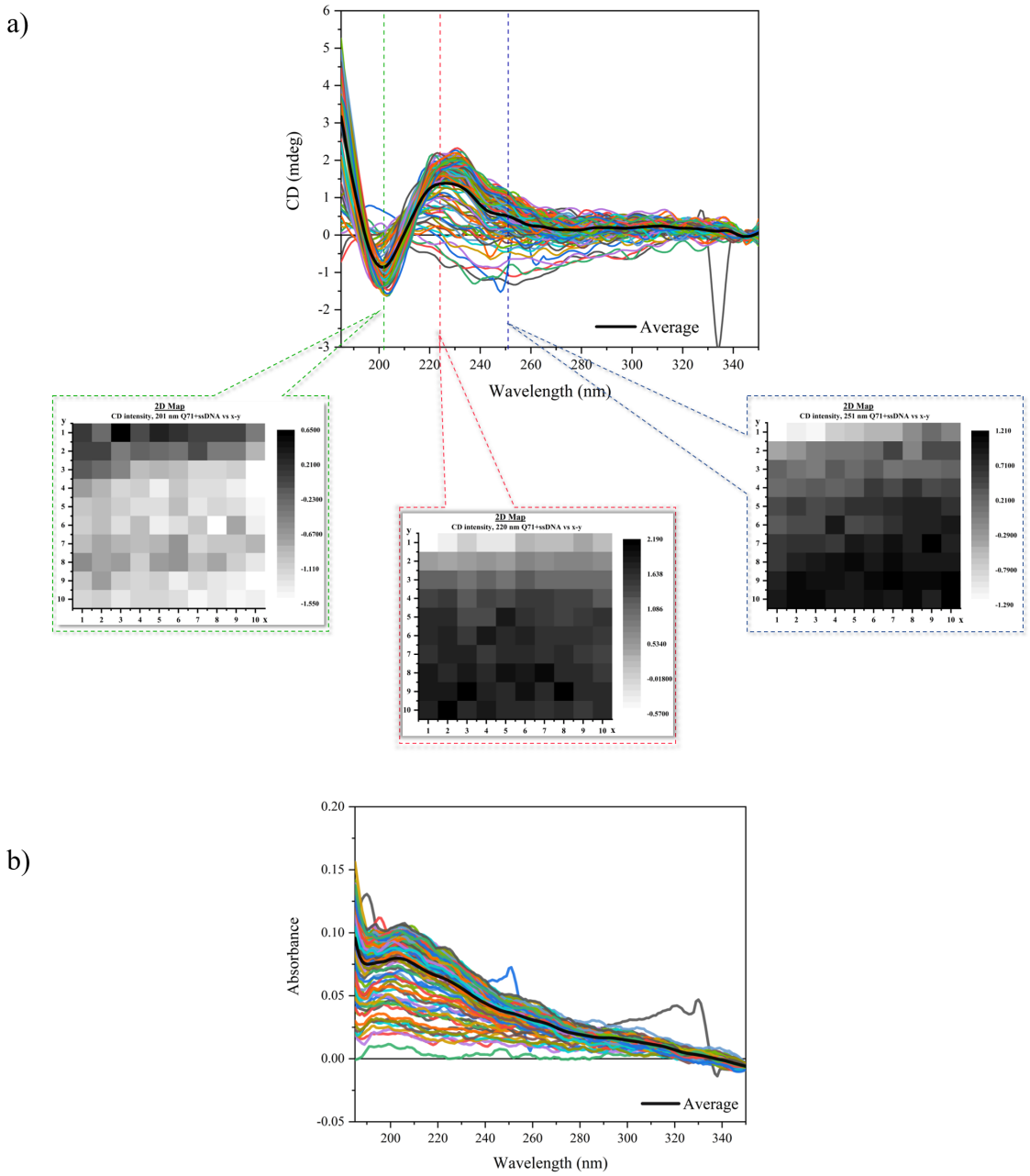

Figure 4.13: a) CD and **b)** absorbance data generated by DX-array subtraction from the DX-array with Q71 attached sample. 2D maps of SRCD intensity at 202, 220 and 251 nm, respectively, vs x-y; each square represents the spot mapped (60 μm^2) during the 10x10 grid array scan.

The following step was to subtract the DX-array to check if any clear band relevant to the protein would be uncovered. However, it is clear that the peak at ~250 nm has disappeared, and now there is a strong positive band at 220 nm and a

negative at 200 nm (Figure 4.13 (a)). This result looks like a random signature, which lead to uncertainty reading of the protein structure upon hybridisation to the DNA nanostructure, or intensity of the signal.

The same change in the shape of the signal is also noticed at the UV absorbance data (Figures 4.12 and 4.13 (b)). In general, however, the absorbance is not at an adequate level. Some visible spikes are due to the beam loss while the experiment was running.

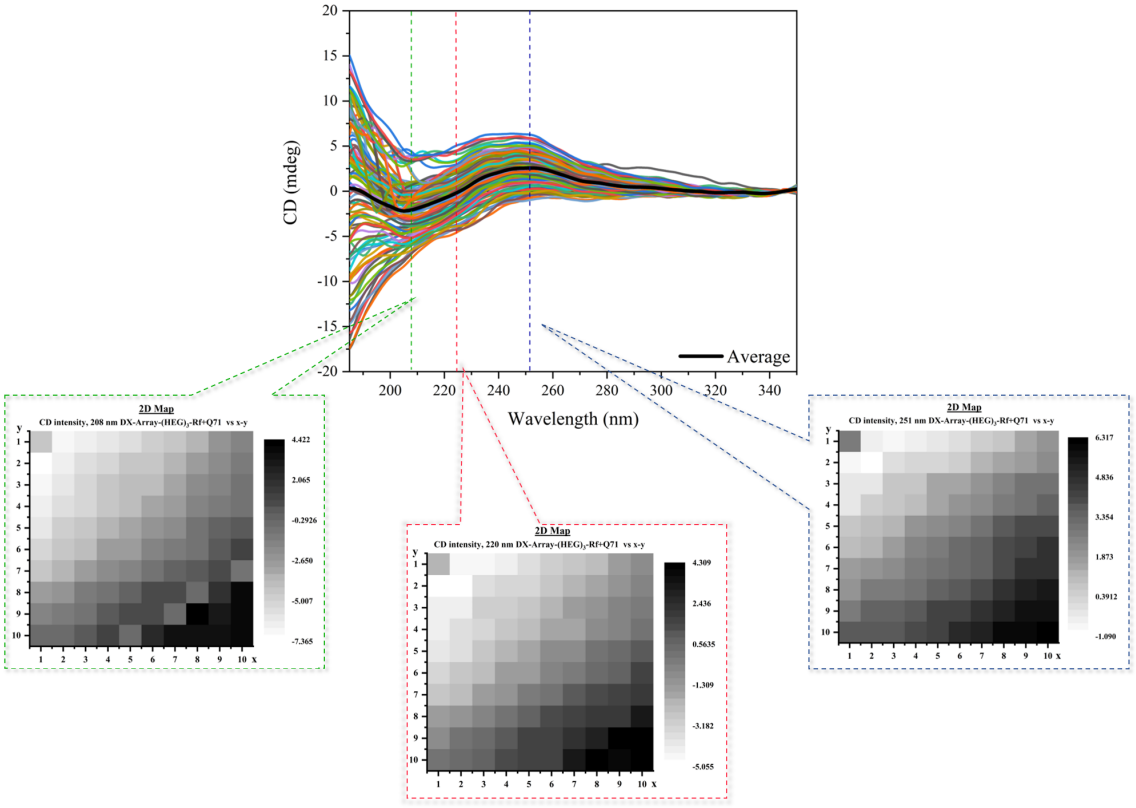

Figure 4.14: CD data of DX-array with Q71 attached as they were recorded during CD*i* experiment by scanning 10x10 grid array of 100 μm using the crucible modified with a big squared pattern of fluorous. 2D maps of SRCD intensity at 208, 220 and 251 nm, respectively, vs x-y; each square represents the spot mapped (100 μm^2) during the 10x10 grid array scan.

The DX-array with Q71 was finally tested in the crucible with the modified single squared area, performing a CD*i* of a 10x10 grid area with 100 μm step. According to Figure 4.14, the CD signal is very similar the results of the crucible with the mini squares. The CD is much stronger here, since some of the scans yield CD above 2 mdeg. A positive peak at 251 nm is present which belongs to the DNA and some very weak minima are also observed at around 222 nm and 208 nm which could result from the presence of the protein, however this is not conclusive. The 2D colour maps plotted using the data from 208, 220 and 251 nm exhibit a comparable pattern and it is noticeable that the sample is not evenly bound on the surface, since the CD is stronger in some mapped areas. They all show a

decreasing intensity from bottom right to top left. This seems to be a systematic behaviour or error, but it cannot be determined whether it occurs due to the sample preparation or the data recording.

The data obtained from this experiment were also corrected against the data of the DX-array itself which were collected earlier using the same crucible, in order to extract the data solely for the protein (Figure 4.15). These spectra did not resemble the α-helical protein, even though a strong minimum at 208 nm is present for some scans. This could be due to the subtraction, or it is an indication that is not an actual characteristic peak of the protein but a random shoulder of the signal recorded. The 2D maps show very similar behaviour with the ones before the DX-array subtraction.

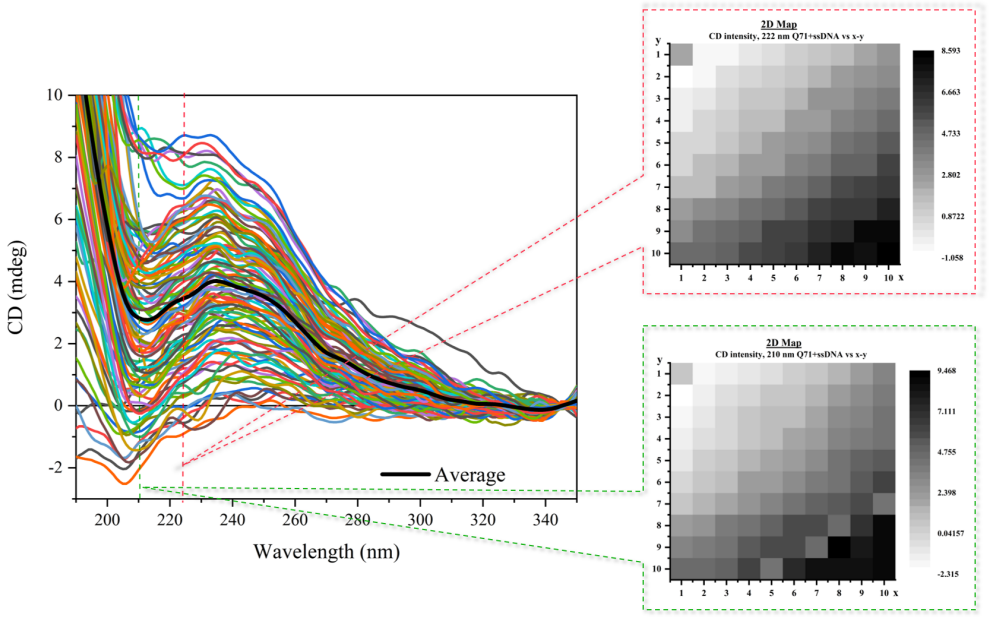

Figure 4.15: CD data generated by DX-array subtraction from the DX-array with Q71 attached sample using the crucible with the mini squares. 2D maps of SRCD intensity at 202, 220 and 251 nm, respectively, vs x-y; each square represents the spot mapped (100 μm^2) during the 10x10 grid array scan

DX-array with Q50 cytb_{562} variant

The final sample that tested using both crucibles was the DX-array hybridised with the long axis variant of cytb_{562} Q50 (dimer). A 10x10 grid array was mapped with 100 μm step size. The graphs in Figure 4.16 show the data recorded using the crucible with the mini squares after being corrected against the baseline (empty crucible). The scans do not show any particular CD signature for either the DNA or the protein (Figure 4.16). Even after the DX-array subtraction (Figure 4.17), the signal looks random even though it is very consistent across the grid array mapped.

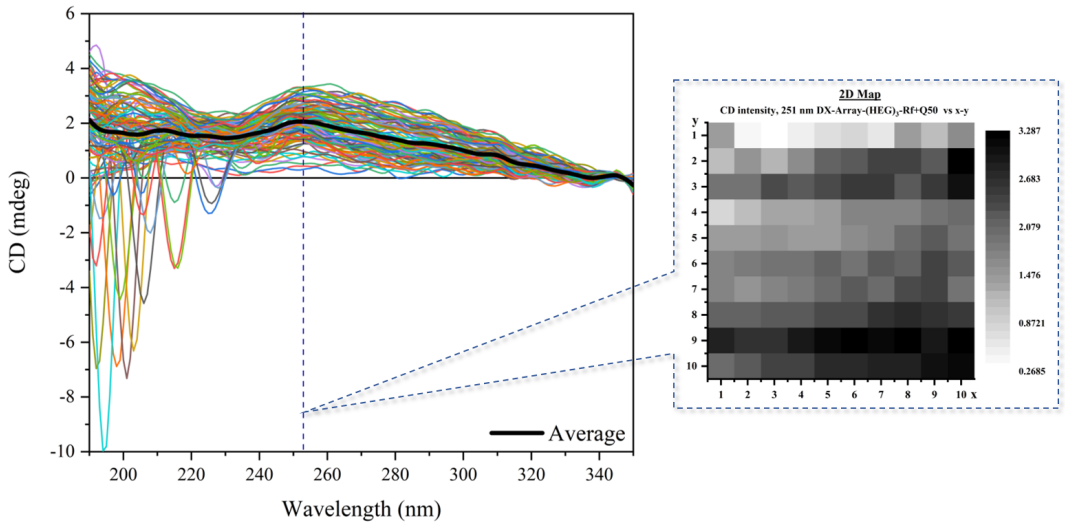

Figure 4.16: CD data of DX-array with Q50 attached as they were recorded during CD*i* experiment by scanning 10x10 grid array of 100 μm using the crucible with the mini perfluorinated squares. 2D map of SRCD intensity at 251 nm, respectively, vs x-y; each square represents the spot mapped (1000 μm^2) during the 10x10 grid array scan.

In the final experiment, the same sample was used but this time it was immobilised on the crucible with the big perfluorinated square. The data recorded and baseline corrected, as well as the data generated after the DX-array subtraction (Figures 4.18 and 4.19, respectively) show comparable results with the Q71 when the same crucible was used (Figures 4.14 and 4.15). The 2D maps are also similar to all the previous results obtained with the same crucible.

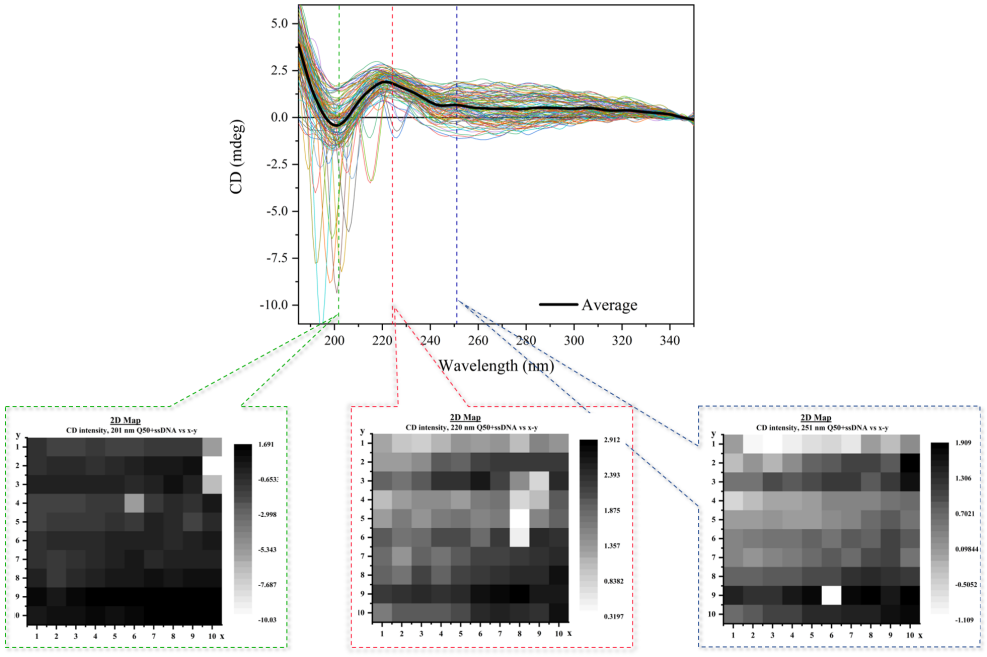

Figure 4.17: CD data generated by DX-array subtraction from the DX-array with Q71 attached sample using the crucible with the mini squares. 2D maps of SRCD intensity at 202, 220 and 251 nm, respectively, vs x-y; each square represents the spot mapped (1000 μm^2) during the 10x10 grid array scan.

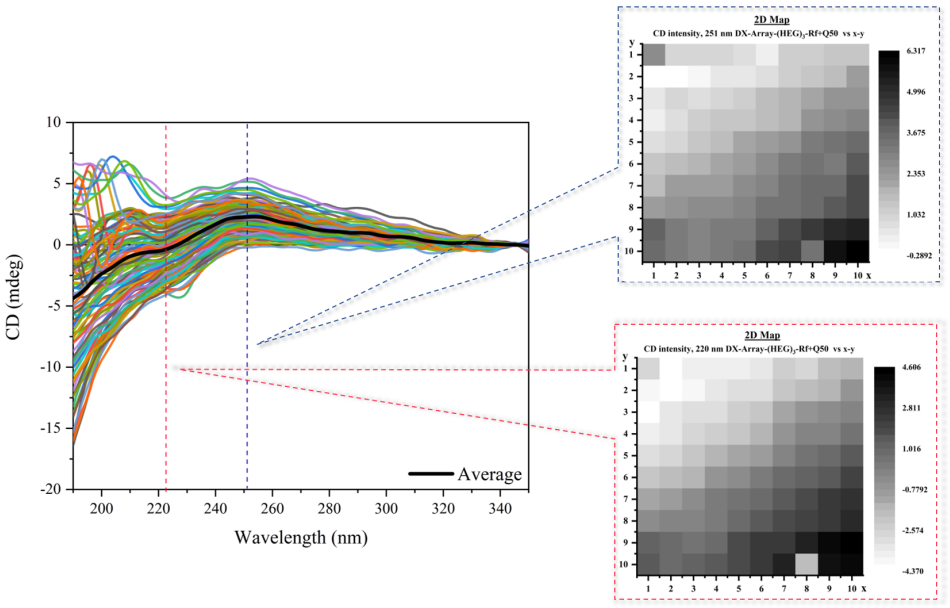

Figure 4.18: CD data of DX-array with Q50 as recorded during CD*i* experiment by scanning 10x10 grid array of 100 μm using the crucible modified with a single square of fluorous modifications. 2D map of SRCD intensity at 220 and 251nm, respectively; each square represents the spot mapped during the 10x10 grid array scan.

4.3 Discussion and conclusion

In this Chapter, a new approach for immobilisation of DNA nanostructures onto quartz crucibles was applied, since various challenges were faced with the film creation in the previous two Chapters. This new approach was based on the fluorous effect which can be used for the immobilisation of fluorous-tagged DNA onto fluorinated-solid supports. This method of attachment allows the absorption of the DNA nanostructure in a more controllable way, the surface can be washed in order to remove any remaining salt, unbound DNA strands and proteins, but, most importantly, the immobilisation is fully reversible, enabling regeneration and re-patterning of the surface with a new fluorous-tagged DNA system.

In order to adapt the fluorous effect method in the system of the DX-array nanostructure, the DX-B tile, which generates along with the DX-A tile the DX-array, was modified with fluorous ponytails with different fluorous content, which included different linkers between the fluorous chain and the carbohydrates. It was shown that regardless of the number of fluorous chains included in the tile and the linkers available, the individual tile and the DX-array were formed successfully. Successful was also the immobilisation of the DX-array onto the perfluorinated quartz crucible confirmed by fluorescence microscopy where GFP protein was used in order to visualise the system. Two fluorinated fused silica crucibles were fluorinated, with the first fluorous array including multiple 50x50 μm mini squares as a sensing area extended all along the surface of the substrate, whereas the second

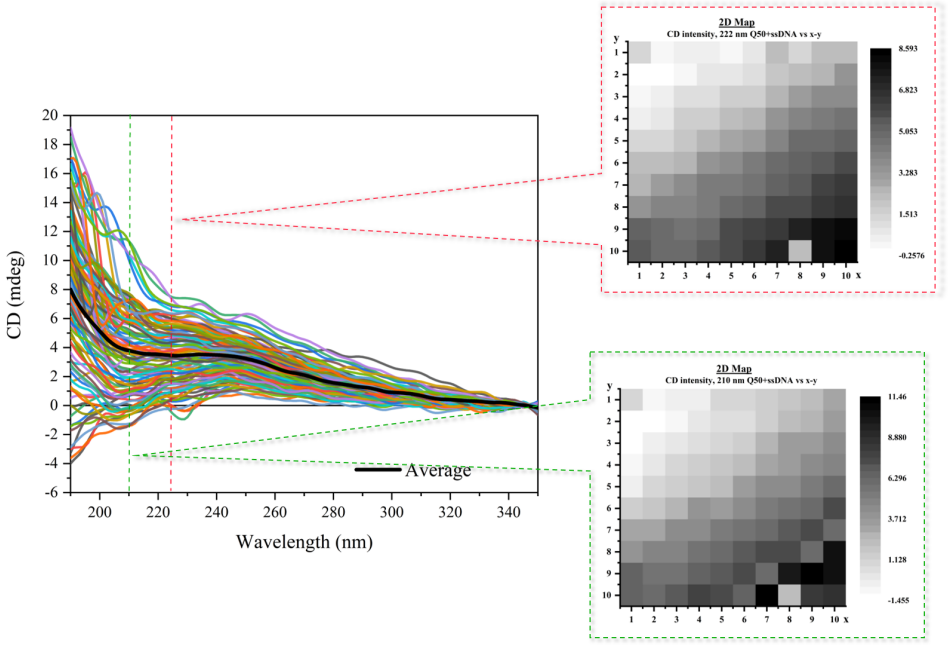

Figure 4.19: CD data generated by DX-array subtraction from the DX-array with Q50 attached sample using the crucible with a big square modified with fluorous. 2D maps of SRCD intensity at 210 and 222nm, respectively; each square represents the spot mapped during the 10x10 grid array scan.

one with a single large square of 3x3 mm and finally, both crucibles were used for SRCD experiments.

The approach regarding the attachment of the DX-array to quartz surfaces via fluorous interactions and then the study of the system before and after protein hybridisation using CD imaging seemed very promising, since, according to the fluorescence microscopy, the DX-array can be immobilised to the quartz crucibles successfully. The challenging part comes with the CD experiments that are not so conclusive. Therefore, the system needs to be further investigated in order to offer better quality in the CD experiments and ensure whether the spectra reflect the DNA-protein samples or just the birefringence of the fused silica crucible.

Chapter 5

Conclusions & Future Work

The DNA origami tiles have been evolved as extremely versatile templates for assembling proteins. It was indicated in previous Chapters that precise orientation of the proteins on origami surfaces is possible through well-defined protein-DNA conjugates, achieved via click chemistry with engineered proteins or enzymes. This project made a step forward by adapting this technology and transferring it to surfaces in order to establish a new protocol for studying precise oriented proteins on DNA nanostructures using SRCD spectroscopy at Diamond Light Source B23 beamline.

In order to carry through with the development of this new technology, different DNA based long-range templates were investigated, which form extremely long structures. They were intended to be utilised as templates for fine-tuned hybridisation of the engineered proteins. The first DNA nanostructure examined was the 2D DNA nanogrid which proved to be very challenging, since despite its successful formation, there were secondary structures and unbound tiles or DNA strands in the sample that could not be removed, and therefore they could have negatively affect the process. The GFP protein was used to be hybridised with an ssDNA via click chemistry, in order to hybridise to the DNA nanostructure. Upon AFM analysis it was shown that it can be successfully be introduced to the DNA nanogrid.

The DNA nanogrid system before and after protein attachment was further analysed with SRCD spectroscopy in films. The films were first tested using conventional benchtop CD in order to ensure the good quality of the films and by, extension, that the recording of either DNA or GFP was possible. The CD data for the DNA nanogrid showed the characteristic peaks of the DNA. However, the size of the beam in benchtop CD instruments is large and as a result, the data

collected represent the average of a large area of the film. This can be misleading, since important conformational information may be missing which is crucial especially for the studies of oriented samples on surfaces. On the other hand, the CD*i* method, which is widely used at B23 beamline at Diamond Light Source, is different compared to the conventional CD methods since it is exploiting the photon flux microbeam of the synchrotron, along with the X-Y stage (available only in B23 beamline) in order to probe smaller and different areas of the film. That gives the opportunity to avoid the non homogeneous areas of the film or areas that could lead to artefacts. Furthermore, structural changes can be spotted easily. The only challenge is the preparation of appropriate samples that are homogeneous in terms of thickness and concentration. In the case of the standalone 2D DNA nanogrid, the characteristic peaks were also observed after mapping an area of 280 μm x 280 μm using the CD*i* method. However, the CD signal was very weak and after the analysis of the data with respect to the g-factor, the film proved to be inhomogeneous in terms of concentration and structure. After visual inspection, the coffee ring effect was observed since various sample layers were present, as well as partial crystallisation especially at the edges of the sample. In the case of the 2D DNA nanogrid with GFP attached, unfortunately, neither of the components provided CD response.

In order to overcome the majority of the difficulties faced with the 2D DNA nanogrid and GFP, a different system was introduced for investigation, the DX-array, along with a new engineered protein, the cytb_{562}. In contrast to the 2D DNA nanogrid that included an anchoring point for protein hybridisation, the original DX-array did not offer one. Therefore, the original design was modified and in the end it provided a protruding area for protein immobilisation. The DX-array was easier to form, yielding the expected pattern. The hybridisation of the different variants of the protein was also successful. CD experiments were performed both in solution and in films. For the experiments in solution, both the DX-array and protein variants even before of after their hybridisation to the DNA nanostructure showed the characteristic CD signatures. The different variants of the protein showed diverse CD intensity depending on which residue the protein was engineered, as well as the position they were immobilised on the DNA nanostructure (either closer to the DNA surface or far from it). On the other hand, in the experiments with films, the preparation of homogeneous DX-array samples was successfully determined using the CD*i* method. The protein CD data showed variations among the different areas scanned indicating that the films were not homogeneous. Visual inspection also confirmed regions of aggregation or crystallisation of the sample which can lead to artefacts.

Even though the efforts presented in this thesis have started to standardise the

preparation methods of the films, since the preparation of uniform films (in terms of concentration and thickness, without the "coffee ring" effect) was successful, the consistent formation of homogeneous films is still unresolved and most of the times visual inspection is not enough to confirm the homogeneity of the films. This can be then confirmed by the CD*i* method. However, it is important not to focusing the experiments on samples that are not suitable for the SRCD experiments. Therefore, B23 beamline will allocate a 3D optical profilometer in the future which can provide information about the thickness and uniformity of the film samples prior to SRCD experiments. This method can be applied as an intermediate step in order to ensure high quality of the films prepared for the CD*i* experiments, and as a result, higher quality data will be collected. Another method for deposition of the samples on the fused silica surfaces which could be explored in the future is the micro bubble spraying. This method is very fast with devoid of any coffee ring formation. Other methods for deposition of the samples on the fused silica surfaces that could be explored in the future are the micro bubble spraying, which is a very fast method with devoid of any coffee ring formation, and the solvents with lower volatility. The challenge with the latter is that there is no vacuum chamber at hand on the beamline, and it would lead to issues with blow-drying of the samples. However, this would be tested by using systems outside of the beamline.

Moreover, another approach for immobilisation of DNA nanostructures onto fused silica crucible was adapted and investigated which was based on the fluorous effect. A fluorinated crucible and a fluorous-tagged DNA were required for the execution of this approach. The system can be absorbed in a more controllable way, the surface can be washed in order to remove unwanted and unbound components and most importantly, the immobilisation is fully reversible, allowing regeneration and re-patterning of the surface with a new fluorous-tagged DNA system. Since the DX-array is the DNA nanostructure that forms easily, it was chosen for the investigation of this approach. Different number of fluorous ponytails and different linkers between the fluorous chain and the carbohydrates were tested and regardless of the number of fluorous chains included in the DNA nanostructure, the DX-array was formed successfully. Furthermore, the successful immobilisation of the fluorous-tagged DX-array onto the fluorinated quartz crucible was confirmed by fluorescence microscopy since the DX-array was hybridised with the GFP protein for the completion of this analysis. Nevertheless, the CD*i* experiments were challenging as the CD signal of either the DX-array or the protein seemed nonexistent. In some CD*i* measurements of the standalone DX-array, it seemed that there could be some signal coming from the DNA but it could also well be that the birefringence of the fused silica was characterised. Unfortunately, the beamtime session in which this approach was investigated for the first time, it was the

last one for this PhD studentship and therefore, it was not explored further. This new immobilisation approach though requires more optimisation in order to ensure that it is suitable for the SRCD and CD*i* measurements. For further exploration of this method, different number of fluorous ponytails included on the DNA nanostructures can be explored in order to find the system that can be immobilised more stably in the crucibles. Finally, Fluorescence Detected Circular Dichroism (FD-CD) can shed more light in whether the fluorous effect as an immobilisation approach can be used with SRCD experiments.[155,156] If this is the case, different proteins can be studied as a proof of concept in order to further develop this technology for protein studies described in this PhD thesis.

Chapter 6

General Experimental Procedures

6.1 Materials & Instruments

The phosphoramides and all the reagents (acetonitrile, deblocking reagent, DCM, activator, capping A & B, oxidizer) were purchased from Tides Service Technology, Germany or Sigma-Aldrich (Merck). The Controlled Pore Glass (CPG) columns for the synthesiser were procured from Sigma-Aldrich (Merck).

The deprotection of the strands was carried out in aqueous concentrated ammonia (Fischer Scientific, UK). Glen-Pak columns were supplied from Glen Research. Amicon filters were purchased by Merck Millipore. Milli-Q water was collected using the Q-Gard purification filter, connected to the Milli-Q Gradient A10 system of Merck Millipore.

For the annealing of the 2D DNA nanostructures, T100 Thermal Cycler from Bio Rad was used. Atomic Force Microscopy (AFM) images under dry conditions were collected using Nanoscope III controller in Southampton University and AFM images in liquid were collected using the Dimension Icon ScanAsyst AFM from Bruker (USA) at Strathclyde University in Glasgow. Mica substrate was also used for the imaging and it was purchased from Agar Scientific. The tips used were supplied by NanoWorld.

UltraViolet-Visible spectroscopy (UV-Vis) for the optical characterization and the measurement of the absorbance of the DNA strands was performed using a Varian Cary 300 Bio UV-Visible spectrophotometer with quartz cells of 1 cm path length, supplied by Hellma.

For the measurement of GFP emission, the Varian Cary-Eclipse Fluorescence spectrophotometer was used with quartz cells of 1 cm path length, supplied by Hellma.

The excitation wavelength for fluorescence spectroscopy were the λmax values as found by UV-Vis.

CD data were obtained using the ChirascanPlus CD instruments and SRCD data were measured using the Module B (Linkam stage and) at B23 beamline at Diamond Light Source, UK. The quartz crucible used for the sample preparation was purchased from Linkam Scientific.

6.2 DNA Synthesis

The 392 DNA/RNA Synthesiser and the Expedite Synthesiser, both from Applied Biosystems, were used for the synthesis of the strands of the 2D nanogrid and the DX-array, respectively. The DNA strands were synthesised by solid phase synthesis using 1000 Å pore CPG beads on a 1 μmol scale. Standard coupling times were used for the bases.

The oligonucleotides were cleaved from the solid support by pushing 1 mL of aqueous ammonia through the columns for 1 hour at room temperature. The bases were deprotected in aqueous ammonia solution by either heating to 55 °C for 4 hours or 40 °C overnight using an Eppendorf Thermomixer Compact. Finally, the oligonucleotides were purified by Glen Pak cartridge column chromatography.

6.3 Glen Pak Cartridge Purification

After the protecting groups have been cleaved for the bases, the oligonucleotides were purified using a Glen Pak column. If this type of purification is to be used, the DNA strands are synthesised DMT-ON. This involves a reverse phase (RP) column, which relies upon binding of the 5'- DMT group to the stationary phase. This allows the failure sequences to be washed through the column.

The sample was first prepared by adding 100 mg/mL Sodium Chloride solution (1 mL). The cartridge was prepared by washing with Acetonitrile (0.5 mL) followed by 2M Triethylamine acetate (TEAA) (1 mL). The acetonitrile washes away organic residues from the resin, while the TEAA behaves as an ion-pairing reagent in order to enhance the binding of the DNA to the resin.

The DNA salt solution was applied to the cartridge and washed with Salt Wash solution (5% acetonitrile in 100 mg/mL sodium chloride) (2 mL). The wash solution rinses away the remaining failure sequences. The cartridge was then washed with 2% Trifluoroacetic acid (TFA) (2 mL), which cleaves the DMT from the bound,

full-length oligonucleotide, directly followed by deionized water (2 mL) to rinse away the excess salt. The sample is then eluted using a 50% acetonitrile in water containing 0.5% ammonium hydroxide (1 mL).

The solvent was removed from the DNA sample using an Eppendorf Concentrator 5301 (speed-vac) at room temperature. The selected settings for the evaporation were based on the solvent used. The resulting solid DNA re-dissolved in deionised water for further analysis.

6.4 High Performance Liquid Chromatography

The purity of the oligonucleotides was further analysed using a Varian 920-LC HPLC and the samples were run in a Varian Polaris 3 C18-A column (150 x 4.6 mm).

5 μL of each strand was injected in 0.1 M TEAA buffer. Depending on the number of bases included in the sequence, two different methods were created. For the long strands (up to 100 bases), the method 1 was used (Table 6.1) and for short strands (up to 25 bases), the method 2 (Table 6.2). In both methods, the column was heated up to 60 °C for better separation. Flow rates were set to 0.6 mL min-1.

Time / min	**A% TEAA Buffer**	**B% MeOH (HPLC grade)**
0	100	0
40	0	100
50	0	100
60	100	0
65	100	0

Table 6.1: Description of **method 1** and solvents used for long DNA strands.

Time/min	**A% TEAA Buffer**	**B% MeOH (HPLC grade)**
0	100	0
40	60	40
50	0	100
60	0	100
70	100	0

Table 6.2: Description of **method 2** and solvents used for short DNA strands.

6.5 Denaturing Urea Gel-PAGE (DNA Purification)

The denaturing urea polyacrylamide gel was first made using the amounts shown in the following Table 6.3.

	Denaturing urea polyacrylamide gel (12.5 %)
UreaGel 29:1 Concentrate (National Diagnostics)	12.5 mL
Urea in deionised aqueous solution, 8 M	10.0 mL
10 x TBE buffer (pH 8.3) with urea, 8 M	2.5 mL
TEMED	20 μL
10% APS	100 μL

Table 6.3: Required amounts of reagents for the preparation of denaturing urea polyacrylamide gel.

All the reagents shown in Table 6.3 were mixed together and the gel solution was added to the frame. Comb was inserted and the gel was allowed to polymerise for one to two hours. After one to two hours of polymerisation, the comb was removed and the gel was placed in the running apparatus. The upper and lower chamber were filled with 0.5 x TBE buffer (0.089 M Tris base, 0.089 M boric acid pH 8.3 and 2 Mm EDTA). DNA was loaded onto the gel using 90 μL of DNA solution and 10 μL of DI stop solution (blue dye with 80% formamide, 10mM EDTA, 5 mM NaOH and 0.1% bromophenol blue). The gel was run for 1 hour in 800 V.

The gel was then separated from the plates and imaged using a fluorescent indicator (UV lamp). The main band was cut out and the DNA was recovered using a "crush and soak" method. The band of the gel used was crushed and transferred to a 1000 μL pipette tip in a microfuge tube. 400 μL of TE solution (10 mM Tris and 0.1 mM EDTA) was added and the tube with tip in was placed in a rotary shaker in order to recover the DNA overnight at room temperature.

Ethanol precipitation was followed for concentrating and de-salting the DNA. The tube with the DNA was spinned in order to collect the entire liquid sample in the tube (the gel remained in the pipette tip). 100% Ethanol (3 x the volume of the DNA sample) and 50 μL of Sodium Acetate (3 M) were added in the tube, which was then placed in dry ice. The DNA was chilled for at least 30 minutes. Following, the DNA was spinned for 20 minutes using centrifuge. The supernatant was removed and the tube with the precipitated DNA was placed in the speed-vac for evaporation from the excess ethanol. In the end, the DNA was dissolved in 100 μL of deionised water in order to measure the absorbance.

6.6 DNA modification with 5' BCN-CEP II modifier

The synthesis of the DNA strand and its modification were carried out on an Applied Biosystems Expedite machine. The modifier, 5'-Click-easy™ BCN CEP II dissolved in 1.5 mL of acetonitrile and it was placed in the modifiers' position at the Expedite. Extended coupling time up to 8 minutes was used for the modification of the oligonucleotide. The modifier was clicked at the 5'-end of the strand. The BCN-DNA was cleaved from the beads using 1.5 mL aqueous ammonia and it was then left for deprotection in the thermomixer at 40 °C overnight. The DMT-OFF, deprotected oligo was evaporated to dryness using the speed-vac and, subsequently, desalting procedure was performed using Glen-Pak.

6.7 Desalting Procedure of BCN-DNA (Glen-Pak)

The fully deprotected, DMT-OFF and dried BCN-DNA was reconstituted in 2.0 mL 0.1M TEAA. In the meantime, the cartridge was prepared using 1.0 mL of Acetonitrile, followed by 1.0 mL of TEAA. The BCN-DNA solution was loaded onto the Glen-Pak column, which was then flashed with 3 x 1.0 mL of deionised water. This wash removes the excess salt from the cartridge. The desalted BCN-DNA was eluted by flashing the cartridge with 1.0 mL 50% Acetonitrile in water. In, the end the sample was evaporated using the speed-vac and it re-dissolved in deionised water for further analysis.

6.8 Protein-DNA modification (Click Chemistry)

sfGFP (1 equiv) was mixed with modified 5'-BCN DNA (3 equiv) in 100 mM sodium phosphate buffer (PBS, pH 8.0, reaction volume 250 μL) for 48 hours at room temperature in the dark. The progress of the modification was monitored using SDS-PAGE. The mixture was concentrated using Amicon filters (10 kDa). The PBS buffer was prepared using the following amounts (Table 6.4).

6.9 SDS-PAGE

A resolving and a stacking gel buffer (stock solutions) were first made using the following amounts shown in the Table 6.5.

	PBS (pH 8.0)
Sodium phosphate monobasic solution, 0.2 M	1.325 mL
Sodium phosphate dibasic solution, 0.2 M	23.675 mL
Sodium Chloride	0.29 g
Deionised water	Up to 50 mL

Table 6.4: Required amounts of reagents for the preparation of PBS buffer.

	Resolving gel buffer (4X, pH 8.8)	Stacking gel buffer (4X, pH 6.8)
Trizma® base, 1.5 M	9.1 g	3.03 g
10% SDS	10.0 mL	0.2 g
TEMED	2.5 mL	0.2 g
Deionised water	Up to 50 mL	Up to 50 mL
The pH was adjusted using HCl.		

Table 6.5: Required amounts of reagents for resolving and stacking gel buffers for SDS-PAGE.

Those buffers were used for the preparation of the resolving and the stacking gel, respectively, mixed with other reagents. The required amounts are shown in the Table 6.6.

	Resolving gel	Stacking gel
Deionised water	3.2 mL	1.55 mL
30% Acrylamide gel solution	4.2 mL	0.33 mL
Resolving gel buffer (4X)	2.5 mL	-
Stacking gel buffer (4X)	-	0.62 mL
10% APS	0.1 mL	0.015 mL

Table 6.6: Required amounts of reagents for resolving and stacking gel for SDS-PAGE.

The loading and the running buffer were also prepared using the following amounts shown in the Tables 6.7 & 6.8, respectively.

The resolving gel (Table 6.6) was first mixed, poured into the mould and left to set for 30 minutes. Isopropyl alcohol (IPA) was used in order to ensure an even gel and then it was poured out. The stacking gel (Table 4.6) was mixed and poured on top of the resolving gel. A comb was inserted to form the wells to load in the protein samples. The gel was left to set for further 15-20 minutes and then the comb was removed. In the meantime, the samples were prepared by incubating them in 2 x loading buffer (Table 6.7) at 95 °C for 15 minutes and then, they were

	2 x loading buffer
Tris HCl, 100 mM	2.5 mL
4% w/v SDS	5.0 mL
0.05% Bromophenol Blue	1.25 mL
20% Glycerol	10.0 mL
DTT, 50 mM	1.25 mL
Deionised water	Up to 25 mL

Table 6.7: Required amounts of reagents for 2 x loading buffer for SDS-PAGE.

	10 x running buffer
Trizma® base, 0.12 M (pH 8.3)	30.0 g
Glysine, 1.25 M	144 g
0.5% w/v SDS	10 g
Deionised water	Up to 1 L

Table 6.8: Required amounts of reagents for 10 x running buffer for SDS-PAGE.

loaded onto the gel. The gel was run at 150-200 V for 45 minutes in 1 x running buffer (Table 6.8). In the end, the gel was stained using InstantBlue™ Coomassie protein stain.

6.10 DNA nanostructures' annealing protocols in solution

The annealing of the oligonucleotides for the formation of the individual tiles and the 2D nanostructures was carried out with two different protocols. In both annealing protocols, the buffer, which was used, was 1 x TAE/Mg^{2+}. However, 10 x TAE/Mg^{2+} buffer was prepared, using the amounts shown in Table 6.9.

	10 x TAE/Mg^{2+} pH 8.0
Trizma® base, 400 mM	4.8 g
Acetic Acid, 200 mM	1.15 mL
EDTA, 20 mM	0.58 g
Magnesium Acetate, 125 mM	2.7 g
Deionised water	Up to 100 mL
The pH was adjusted using acetic acid.	

Table 6.9: Required amounts of reagents for 10 x TAE/Mg^{2+} buffer for annealing.

Protocol F1: All the single strands involved in the tiles were mixed in an Eppendorf tube with 1 x TAE/ Mg^{2+} buffer in total volume of 60 μL and concentration 1 μM. The mixture was heated up to 90 °C and cooled down to 4 °C in a period

of $\sim$ 13 hours on T100 Thermal Cycler. The whole protocol used is shown in the Table 6.10.

Step	Temperature, °C	Hold duration/ Speed of cooling
1	90	2 minutes and 30 seconds
2	90 → 70	0.2 °C / 8 seconds
3	70 → 60	0.1 °C / 8 seconds
4	60 → 50	0.1 °C / 25 seconds
5	50 → 45	0.1 °C / 50 seconds
6	45 → 37	0.1 °C / 2 minutes
7	37 → 20	0.1 °C / 3 minutes
8	20 → 16	10 minutes
9	4	∞

Table 6.10: Annealing protocol F1 (90 °C → 4°C).

Protocol F2: Each individual tiles were assembled by mixing a stoichiometric quantity of the DNA strands involved in the tile in 1 x TAE/Mg^{2+} buffer. The oligo mixtures were cooled from 90 °C to 4 °C in a period of $\sim$ 13h using the T100 Thermal Cycler (Table 6.10). Then, a ratio of 1:1 of each tile was mixed, heated up to 50 °C and cooled to 4 °C in a period of $\sim$ 6 hours, in order to form the 2D DNA nanostructures (Table 6.11).

Step	Temperature, °C	Hold duration/ Speed of cooling
1	50	10 minutes
2	50 → 40	0.2 °C / minute
3	40	1 hour
4	40 → 35	0.2 °C / minute
5	35 → 30	0.2 °C / minute
6	30 → 25	0.2 °C / minute
7	25 → 20	0.2 °C / minute
8	20 → 15	0.2 °C / minute
9	15 → 10	0.2 °C / minute
10	10 → 5	0.2 °C / minute
11	5	1 minute
12	4	∞

Table 6.11: Annealing protocol F2, second half 50 °C → 4 °C.

Formation of the DNA tiles and the 2D DNA nanostructures was analysed by 1% Agarose gel electrophoresis and native PAGE.

6.11 Agarose Gel 1%

1 x TAE / Mg^{2+} was first prepared using the amounts shown in the Table 6.12. This buffer was used both for the agarose gel preparation and as a running buffer.

	10 x TAE/Mg^{2+} pH 8.0
Trizma® base, 40 mM	4.8 g
Acetic Acid, 20 mM	1.15 mL
EDTA, 2 mM	0.58 g
Magnesium Acetate, 11 mM	2.4 g
Deionised water	Up to 1 L
The pH was adjusted using acetic acid.	

Table 6.12: Required amounts of reagents for 10 x TAE/Mg^{2+} buffer for annealing.

Agarose (1 g) was suspended in 100 mL 1 x TAE/Mg^{2+} (100 mL) and heated in the microwave until all the solid had dissolved. The mixture was left to cool down for 5 minutes. SYBR gold (10 μL) was then added to the solution before it was poured into a mould and a comb was inserted. The gel was then left to set for 20-30 minutes and, subsequently, the comb was removed. The gel was placed in the running apparatus, which was filled with the 0.5 x TAE/Mg^{2+} buffer, as the gel must be fully covered with the buffer. DNA was loaded in the gel using 2 μL of DNA sample, 8 μL of deionised water and 2.5 μL of 5 x Green GoTaq® buffer. The gel was run for 1-2.5 hours in 60-70 V. Ice bags were used in order to prevent the gel from overheating. In the end, the gel was imaged using the Gel Doc™ EZ imager from BIO RAD.

6.12 Protein-ssDNA Co-Assembly on DNA nanostrustures

The DNA nanostructures in 1 X TAE/Mg^{2+} buffer (1 μM) were mixed with protein conjugated with addressing ssDNA with a molar ratio of 1:1. The solution was left for hybridisation in room temperature over a period of 24 hours. As an alternative, the hybridisation could be carried out from 37 °C to 4 °C over a gradient of -1 °C per minute on T100 Thermal Cycler. The system was then visualised using Atomic Force Microscopy (AFM).

6.13 Sample preparation for AFM

An approximately 1 cm mica disk was attached to a magnetic slide using double-sided conducting adhesive tape. The mica was carefully positioned on the slide and firmly pressed to ensure that no air bubbles formed beneath the disc.

Once the mica was positioned on the magnetic slide, it was cleaved using Scotch® tape to produce an atomically flat surface. The tape was pressed onto the surface and then slowly removed. This removed the very top layer of the mica disk. This was repeated 3-4 times as the first couple of cleaves produced a rather inferior surface.

Sample preparation for the AFM under dry conditions: The sample to be imaged in the appropriate buffer (1X TAE/Mg^{2+}) was pipetted onto the centre of the mica disc (3-5 μL). The solution was left to incubate for 1-3 minutes, depending on the volume of liquid pipetted onto the surface, to ensure that the DNA was completely bound to the mica. It was then washed with 100-200 μL deionised water and dried using a nitrogen gun. The sample was transported to the AFM in a sealed container to prevent any contaminants attaching to the surface and was imaged straight away, using tapping mode. AFM imaging was performed on Nanoscope III SPM controller.

Sample preparation for the liquid AFM: The sample to be imaged in the appropriate buffer was pipetted onto the centre of the mica disc (5 μL) and it was left to incubate for 5 minutes. Then, 20 μL of the annealing buffer was placed onto the AFM tip and on the top of the sample. AFM images were obtained on FastScan Bio™ with ScanAsyst from Bruker (USA) at Strathclyde University in Glasgow. Images were then analysed using NanoScope Analysis 1.5 software.

6.14 CD and SRCD Spectroscopy

The CD experiments were performed using a nitrogen- flushed Chirascan Plus spectropolarimeter (bench-top instrument) from Applied Photophysics (Leathrhead, UK).All the SRCD and CD*i* measurements were recorded at B23, Diamond Light Source, UK, during beamtime allocations. Measurements were recorded using a nitrogen-flushed module B end station which is equipped with a highly modified DSM 20 spectropolarimeter (Olis Inc, Georgia, USA).[96] Prior to the start of the set of experiments, the signal calibration was checked using a standard solution of 0.6 mg mL^{-1} camphor sulphonic acid, which normally gives a ration of 2:1 for the absolute peak intensity at 290.5 nm and 192.5 nm.[157] The samples were measured in the unique B23 vertical sample chamber equipped with motorised XY stage.

The samples were analysed in solution and in solid phase (films). For sample preparation in solution, at least 30 μL of the desired sample was pipetted in a quartz cell of the appropriate path length. The films were prepared by dropping sample (5-15 μL) on a quartz substrate and they were dried under a stream of nitrogen. All the CD and SRCD data were baseline corrected as appropriate and processed further in CDApps.[116]

6.15 Immobilisation of DNA nanostructures on fluorous modified quartz disk

Fluorous-tagged DNA nanostructures were immobilised onto patterned quartz disk with Heptadecafluoro-1, 1, 2, 2-tetrahydrodecyl) trimethoxysilane (FTDS). The solution of the fluorous tagged DNA nanostructure with protein attached (1 μM) formed as described in Sections 6.10 and 6.12 was spotted onto the quartz and left from 30 minutes to 2 hours at room temperature in a humidity chamber (container with a lid, blue roll and parafilm). The quartz surface was then rinsed with 10 μL of water and dried under nitrogen.

An alternative approach of introducing the protein into the DNA nanostructure could be upon the immobilisation of the DNA onto the perfluorinated quartz crucible and when the surface is washed and dried then the protein-ssDNA will be incubated on the quartz following the same procedure.

6.16 Removal of the DNA nanostructure from the fluorous modified quartz surface and its re-immobilisation

The DNA nanostructure was removed from the surface by immersing the substrates in a solution containing 50% MeOH, in TAE/Mg^{2+} buffer, overnight. Following the substrates were removed, rinsed in TAE/Mg^{2+} and water. The re-immobilisation of the DNA nanostructure and hybridisation with the protein-ssDNA was carried out as described in Section 6.15.

References

[1] Stiefler-Jensen, D. Stability of Hyperthermophilic Proteins: Post-translational modification and other factors. Thesis, 2016.

[2] Crick, F. H. On protein synthesis. Symp Soc Exp Biol. p 8.

[3] Pain, R. H. *Mechanisms of Protein Folding*; Frontiers in Molecular Biology (Book 3); Oxford University Press, U.S.A., 1994.

[4] Petsko, G. A.; Ringe, D. *Protein structure and function*; New Science Press, 2004.

[5] Price, N. C.; Nairn, J. *Exploring proteins: a student's guide to experimental skills and methods*; Oxford University Press Oxford, UK, 2009.

[6] Liddington, R.; Derewenda, Z.; Dodson, E.; Hubbard, R.; Dodson, G. High resolution crystal structures and comparisons of T-state deoxyhaemoglobin and two liganded T-state haemoglobins: T(alpha-oxy)haemoglobin and T(met)haemoglobin. *J Mol Biol* **1992**, *228*, 551–79.

[7] David L. Nelson, M. M. C. *Lenhinger Principles of Biochemistry*, 7th ed.; W. H. Freeman: New York, 2008.

[8] Smyth, M.; Martin, J. x Ray crystallography. *Molecular Pathology* **2000**, *53*, 8.

[9] Bragg, W. L.; Kendrew, J. C.; Perutz, M. F. Polypeptide chain configurations in crystalline proteins. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* **1950**, *203*, 321–357.

[10] Carpenter, E. P.; Beis, K.; Cameron, A. D.; Iwata, S. Overcoming the challenges of membrane protein crystallography. *Curr. Opin. Struct. Biol.* **2008**, *18*, 581–586.

[11] Wang, H.; Wang, J. How cryo-electron microscopy and X-ray crystallography complement each other. *Protein Sci.* **2017**, *26*, 32–39.

[12] Park, E.; Campbell, E. B.; MacKinnon, R. Structure of a CLC chloride ion channel by cryo-electron microscopy. *Nature* **2017**, *541*, 500–505.

[13] Penczek, P. A.; Kimmel, M.; Spahn, C. M. Identifying conformational states of macromolecules by eigen-analysis of resampled cryo-EM images. *Structure* **2011**, *19*, 1582–1590.

[14] Carroni, M.; Saibil, H. R. Cryo electron microscopy to determine the structure of macromolecular complexes. *Methods* **2016**, *95*, 78–85.

[15] Duckworth, B. P.; Xu, J.; Taton, T. A.; Guo, A.; Distefano, M. D. Site-specific, covalent attachment of proteins to a solid surface. *Bioconjugate Chem.* **2006**, *17*, 967–974.

[16] Rusmini, F.; Zhong, Z.; Feijen, J. Protein immobilization strategies for protein biochips. *Biomacromolecules* **2007**, *8*, 1775–1789.

[17] Frey, B. L.; Jordan, C. E.; Kornguth, S.; Corn, R. M. Control of the specific adsorption of proteins onto gold surfaces with poly (L-lysine) monolayers. *Anal. Chem.* **1995**, *67*, 4452–4457.

[18] Reed, A. M. W.; Metallo, S. J. Oriented Protein Adsorption to Gold Nanoparticles through a Genetically Encodable Binding Motif. *Langmuir* **2010**, *26*, 18945–18950.

[19] Niu, Y.; Matos, A. I.; Abrantes, L. M.; Viana, A. S.; Jin, G. Antibody Oriented Immobilization on Gold using the Reaction between Carbon Disulfide and Amine Groups and Its Application in Immunosensing. *Langmuir* **2012**, *28*, 17718–17725.

[20] Tang, Y.; Zeng, X.; Liang, J. Surface plasmon resonance: an introduction to a surface spectroscopy technique. *J. Chem. Educ.* **2010**, *87*, 742–746.

[21] Homola, J.; Yee, S. S.; Gauglitz, G. Surface plasmon resonance sensors. *Sens. Actuators, B* **1999**, *54*, 3–15.

[22] Rehman, A.; Zeng, X. Monitoring the Cellular Binding Events with Quartz Crystal Microbalance (QCM) Biosensors. *Methods Mol. Biol.* **2017**, *1572*, 313–326.

[23] Park, S. H.; Yin, P.; Liu, Y.; Reif, J. H.; LaBean, T. H.; Yan, H. Programmable DNA self-assemblies for nanoscale organization of ligands and proteins. *Nano Lett.* **2005**, *5*, 729–733.

[24] Marth, G.; Hartley, A. M.; Reddington, S. C.; Sargisson, L. L.; Parcollet, M.; Dunn, K. E.; Jones, D. D.; Stulz, E. Precision Templated Bottom-Up Multiprotein Nanoassembly through Defined Click Chemistry Linkage to DNA. *ACS Nano* **2017**, *11*, 5003–5010.

[25] Watson, J. D.; Crick, F. H. Genetical implications of the structure of deoxyribonucleic acid. 1953. *Jama* **1993**, *269*, 1967–9.

[26] Pray, L. Discovery of DNA structure and function: Watson and Crick. *Nature Education* **2008**, *1*, 100.

[27] Neidle, S. Principles of nucleic acid structure. **2018**,

[28] Watson, J. D.; Crick, F. H. C. Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature* **1953**, *171*, 737.

[29] Dickerson, R. E.; Drew, H. R.; Conner, B. N.; Wing, R. M.; Fratini, A. V.; Kopka, M. L. The anatomy of a-, b-, and z-dna. *Science* **1982**, *216*, 475–485.

[30] Arnott, S. Historical article: DNA polymorphism and the early history of the double helix. *Trends Biochem. Sci.* **2006**, *31*, 349–354.

[31] 2010; `https://www.mun.ca/biology/scarr/A_B_Z_DNA.html`.

[32] Seeman, N. C. DNA in a material world. *Nature* **2003**, *421*, 427.

[33] Ussery, D. W. DNA Structure: A-, B-and Z-DNA Helix Families. *e LS* **2001**,

[34] Htun, H.; Dahlberg, J. E. Single strands, triple strands, and kinks in H-DNA. *Science* **1988**, *241*, 1791–1796.

[35] Gellert, M.; Lipsett, M. N.; Davies, D. R. Helix formation by guanylic acid. *Proc. Natl. Acad. Sci. U. S. A.* **1962**, *48*, 2013.

[36] Bacolla, A.; Wells, R. D. Non-B DNA conformations, genomic rearrangements, and human disease. *J. Biol. Chem.* **2004**, *279*, 47411–47414.

[37] Pandey, S.; Ogloblina, A. M.; Belotserkovskii, B. P.; Dolinnaya, N. G.; Yakubovskaya, M. G.; Mirkin, S. M.; Hanawalt, P. C. Transcription blockage by stable H-DNA analogs in vitro. *Nucleic Acids Res.* **2015**, *43*, 6994–7004.

[38] Rhodes, D.; Lipps, H. J. G-quadruplexes and their regulatory roles in biology. *Nucleic Acids Res.* **2015**, *43*, 8627–8637.

[39] Brown, T. B. T. *Nucleic Acids Book.*

[40] Stulz, E. DNA Architectonics: towards the Next Generation of Bio-inspired Materials. *Chem. Eur. J.* **2012**, *18*, 4456–4469.

[41] Lin, C.; Liu, Y.; Rinker, S.; Yan, H. DNA Tile Based Self-Assembly: Building Complex Nanoarchitectures. *ChemPhysChem* **2006**, *7*, 1641–1647.

[42] Seeman, N. C. Nucleic acid junctions and lattices. *J. Theor. Biol.* **1982**, *99*, 237–247.

[43] LaBean, T. H.; Yan, H.; Kopatsch, J.; Liu, F.; Winfree, E.; Reif, J. H.; Seeman, N. C. Construction, analysis, ligation, and self-assembly of DNA triple crossover complexes. *J. Am. Chem. Soc.* **2000**, *122*, 1848–1860.

[44] Sharma, J.; Chhabra, R.; Liu, Y.; Ke, Y.; Yan, H. DNA-Templated Self-Assembly of Two-Dimensional and Periodical Gold Nanoparticle Arrays. *Angew. Chem., Int. Ed.* **2006**, *45*, 730–735.

[45] He, Y.; Ye, T.; Su, M.; Zhang, C.; Ribbe, A. E.; Jiang, W.; Mao, C. Hierarchical self-assembly of DNA into symmetric supramolecular polyhedra. *Nature* **2008**, *452*, 198.

[46] Yang, Y. R.; Liu, Y.; Yan, H. DNA nanostructures as programmable biomolecular scaffolds. *Bioconjugate Chem.* **2015**, *26*, 1381–1395.

[47] Rothemund, P. W. Design of DNA origami. Proceedings of the 2005 IEEE/ACM International conference on Computer-aided design. pp 471–478.

[48] Yang, Y.; Han, D.; Nangreave, J.; Liu, Y.; Yan, H. DNA origami with double-stranded DNA as a unified scaffold. *ACS nano* **2012**, *6*, 8209–8215.

[49] Endo, M.; Yang, Y.; Sugiyama, H. DNA origami technology for biomaterials applications. *Biomater. Sci.* **2013**, *1*, 347–360.

[50] Douglas, S. M.; Dietz, H.; Liedl, T.; Högberg, B.; Graf, F.; Shih, W. M. Self-assembly of DNA into nanoscale three-dimensional shapes. *Nature* **2009**, *459*, 414–418.

[51] Dietz, H.; Douglas, S. M.; Shih, W. M. Folding DNA into twisted and curved nanoscale shapes. *Science* **2009**, *325*, 725–730.

[52] Zhang, J.; Liu, Y.; Ke, Y.; Yan, H. Periodic Square-Like Gold Nanoparticle Arrays Templated by Self-Assembled 2D DNA Nanogrids on a Surface. *Nano Letters* **2006**, *6*, 248–251.

[53] Niemeyer, C. M.; Koehler, J.; Wuerdemann, C. DNA-directed assembly of bienzymic complexes from in vivo biotinylated NAD(P)H : FMN oxidoreductase and luciferase. *Chembiochem* **2002**, *3*, 242–245.

[54] Fu, J. L.; Liu, M. H.; Liu, Y.; Woodbury, N. W.; Yan, H. Interenzyme Substrate Diffusion for an Enzyme Cascade Organized on Spatially Addressable DNA Nanostructures. *J. Am. Chem. Soc.* **2012**, *134*, 5516–5519.

[55] Selmi, D. N.; Adamson, R. J.; Attrill, H.; Goddard, A. D.; Gilbert, R. J.; Watts, A.; Turberfield, A. J. DNA-templated protein arrays for single-molecule imaging. *Nano Lett.* **2011**, *11*, 657–660.

[56] Saccà, B.; Niemeyer, C. M. Functionalization of DNA nanostructures with proteins. *Chem. Soc. Rev.* **2011**, *40*, 5910–5921.

[57] Van Vught, R.; Pieters, R. J.; Breukink, E. Site-specific functionalization of proteins and their applications to therapeutic antibodies. *Comp. Struct. Biotechnol.* **2014**, *9*, 1–13.

[58] Fu, J.; Yang, Y. R.; Johnson-Buck, A.; Liu, M.; Liu, Y.; Walter, N. G.; Woodbury, N. W.; Yan, H. Multi-enzyme complexes on DNA scaffolds capable of substrate channelling with an artificial swinging arm. *Nat. nanotechnol.* **2014**, *9*, 531–536.

[59] 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. *Nucleic Acids Res.* **1994**, *22*, 5530–5539.

[60] Jones, D. S.; Rowe, C. G.; Chen, B.; Reiter, K.; Rausch, K. M.; Narum, D. L.; Wu, Y.; Duffy, P. E. A Method for producing protein nanoparticles with applications in vaccines. *PloS one* **2016**, *11*, e0138761.

[61] Meredith, G. D.; Wu, H. Y.; Allbritton, N. L. Targeted protein functionalization using His-tags. *Bioconjugate Chem.* **2004**, *15*, 969–982.

[62] Lotze, J.; Reinhardt, U.; Seitz, O.; Beck-Sickinger, A. G. Peptide-tags for site-specific protein labelling in vitro and in vivo. *Molecular BioSystems* **2016**, *12*, 1731–1745.

[63] Yin, J.; Lin, A. J.; Golan, D. E.; Walsh, C. T. Site-specific protein labeling by Sfp phosphopantetheinyl transferase. *Nature protocols* **2006**, *1*, 280.

[64] Niemeyer, C. M. Semisynthetic DNA–protein conjugates for biosensing and nanofabrication. *Angew. Chem., Int. Ed.* **2010**, *49*, 1200–1216.

[65] Lapiene, V.; Kukolka, F.; Kiko, K.; Arndt, A.; Niemeyer, C. M. Conjugation of fluorescent proteins with DNA oligonucleotides. *Bioconjugate Chem.* **2010**, *21*, 921–927.

[66] Duckworth, B. P.; Chen, Y.; Wollack, J. W.; Sham, Y.; Mueller, J. D.; Taton, T. A.; Distefano, M. D. A universal method for the preparation of covalent protein–DNA conjugates for use in creating protein nanostructures. *Angew. Chem.* **2007**, *119*, 8975–8978.

[67] Link, A. J.; Tirrell, D. A. Cell surface labeling of escherichia c oli via copper (i)-catalyzed [3+ 2] cycloaddition. *J. Am. Chem. Soc.* **2003**, *125*, 11164–11165.

[68] Kennedy, D. C.; McKay, C. S.; Legault, M. C.; Danielson, D. C.; Blake, J. A.; Pegoraro, A. F.; Stolow, A.; Mester, Z.; Pezacki, J. P. Cellular consequences of copper complexes used to catalyze bioorthogonal click reactions. *J. Am. Chem. Soc.* **2011**, *133*, 17993–18001.

[69] Reddington, S. C.; Rizkallah, P. J.; Watson, P. D.; Pearson, R.; Tippmann, E. M.; Jones, D. D. Different photochemical events of a genetically encoded phenyl azide define and modulate GFP fluorescence. *Angew.Chem., Int. Ed.* **2013**, *52*, 5974–5977.

[70] Woody, R.; Dunker, A. In Circular Dichroism and the Conformational Analysis of Biomolecules (Fasman, GD, ed.). *Theory of Circular Dichroism of Proteins* **1996**, 25–68.

[71] Kelly, S. M.; Jess, T. J.; Price, N. C. How to study proteins by circular dichroism. *Biochim. Biophys Acta, Proteins and Proteomics* **2005**, *1751*, 119–139.

[72] Cotton, A. Absorption inégale des rayons circulaires droit et gauche dans certains corps actifs. *Compt. Rend* **1895**, *120*, 989–991.

[73] Hussain, R.; Siligardi, G. *The Next Generation in Membrane Protein Structure Determination*; Springer, 2016; pp 43–59.

[74] Fasman, G. D. *Circular dichroism and the conformational analysis of biomolecules*; Springer Science & Business Media, 2013.

[75] Miles, A. J.; Wallace, B. Circular dichroism spectroscopy of membrane proteins. *Chem. Soc. Rev.* **2016**, *45*, 4859–4872.

[76] Whitmore, L.; Wallace, B. A. Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases. *Biopolymers: Original Research on Biomolecules* **2008**, *89*, 392–400.

[77] Hammes, G. G. *Spectroscopy for the biological sciences*; John Wiley & Sons, 2005.

[78] Berova, N.; Nakanishi, K.; Woody, R. W. *Circular dichroism: principles and applications*; John Wiley & Sons, 2000.

[79] Moffitt, W. Optical rotatory dispersion of helical polymers. *The Journal of Chemical Physics* **1956**, *25*, 467–478.

[80] Kelly, S. M.; Jess, T. J.; Price, N. C. How to study proteins by circular dichroism. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics* **2005**, *1751*, 119–139.

[81] Holm, A. I. S.; Nielsen, L. M.; Hoffmann, S. V.; Nielsen, S. B. Vacuum-ultraviolet circular dichroism spectroscopy of DNA: a valuable tool to elucidate topology and electronic coupling in DNA. *Phys. Chem. Chem. Phys.* **2010**, *12*, 9581–9596.

[82] Van Holde, K. E.; Johnson, W. C.; Ho, P. S., et al. Principles of physical biochemistry. **2006**,

[83] Ranjbar, B.; Gill, P. Circular dichroism techniques: biomolecular and nanostructural analyses-a review. *Chem. Biol. Drug Des.* **2009**, *74*, 101–120.

[84] Gray, D. M.; Liu, J.-J.; Ratliff, R. L.; Allen, F. S. Sequence dependence of the circular dichroism of synthetic double-stranded RNAs. *Biopolymers: Original Research on Biomolecules* **1981**, *20*, 1337–1382.

[85] Bloomfield, V.; Crothers, D. M. *Nucleic acids: structures, properties and functions*; 2000.

[86] Sprecher, C. A.; Baase, W. A.; Johnson Jr, W. C. Conformation and circular dichroism of DNA. *Biopolymers: Original Research on Biomolecules* **1979**, *18*, 1009–1019.

[87] Gray, D. M.; Taylor, T. N.; Lang, D. Dehydrated circular DNA: circular dichroism of molecules in ethanolic solutions. *Biopolymers: Original Research on Biomolecules* **1978**, *17*, 145–157.

[88] Riazance, J. H.; Baase, W. A.; Johnson Jr, W. C.; Hall, K.; Cruz, P.; Tinoco Jr, I. Evidence for Z-form RNA by vacuum UV circular dichroism. *Nucleic Acids Res.* **1985**, *13*, 4983–4989.

[89] George, J.; Thomas, K. G. Surface Plasmon Coupled Circular Dichroism of Au Nanoparticles on Peptide Nanotubes. *J. Am. Chem. Soc.* **2010**, *132*, 2502–2503.

[90] Chao, J.; Lin, Y.; Liu, H.; Wang, L.; Fan, C. DNA-based plasmonic nanostructures. *Mater. Today* **2015**, *18*, 326–335.

[91] Schreiber, R. Plasmonic DNA nanostructures with tailored optical response. Thesis, 2014.

[92] Shen, X.; Zhan, P.; Kuzyk, A.; Liu, Q.; Asenjo-Garcia, A.; Zhang, H.; de Abajo, F. J. G.; Govorov, A.; Ding, B.; Liu, N. 3D plasmonic chiral colloids. *Nanoscale* **2014**, *6*, 2077–2081.

[93] Xiaoli, W.; Zhiyong, T. Circular Dichroism Studies on Plasmonic Nanostructures. *Small* **2017**, *13*, 1601115.

[94] Ries, O.; Löffler, P. M.; Vogel, S. Convenient synthesis and application of versatile nucleic acid lipid membrane anchors in the assembly and fusion of liposomes. *Org. Biomol. Chem.* **2015**, *13*, 9673–9680.

[95] Hussain, R.; Jávorfi, T.; Siligardi, G. Spectroscopic analysis: synchrotron radiation circular dichroism in comprehensive chirality. **2012**, *8*, 438–448.

[96] Hussain, R.; Jávorfi, T.; Siligardi, G. Circular dichroism beamline B23 at the Diamond Light Source. *Journal of synchrotron radiation* **2012**, *19*, 132–135.

[97] Bürck, J.; Roth, S.; Windisch, D.; Wadhwani, P.; Moss, D.; Ulrich, A. S. UV-CD12: synchrotron radiation circular dichroism beamline at ANKA. *Journal of synchrotron radiation* **2015**, *22*, 844–852.

[98] Zinna, F.; Resta, C.; Górecki, M.; Pescitelli, G.; Di Bari, L.; Jávorfi, T.; Hussain, R.; Siligardi, G. Circular Dichroism Imaging: Mapping the Local Supramolecular Order in Thin Films of Chiral Functional Polymers. *Macromolecules* **2017**, *50*, 2054–2060.

[99] Siligardi, G.; Hussain, R. CD spectroscopy: an essential tool for quality control of protein folding. *Structural Proteomics: High-Throughput Methods* **2015**, 255–276.

[100] Binnig, G.; Quate, C. F.; Gerber, C. Atomic force microscope. *Phys. Rev. Lett.* **1986**, *56*, 930–933.

[101] Gadegaard, N. Atomic force microscopy in biology: technology and techniques. *Biotech. Histochem.* **2006**, *81*, 87–97.

[102] Tessmer, I.; Kaur, P.; Lin, J. G.; Wang, H. Investigating bioconjugation by atomic force microscopy. *J. Nanobiotechnol.* **2013**, *11*.

[103] Winfree, E.; Liu, F.; Wenzler, L. A.; Seeman, N. C. Design and self-assembly of two-dimensional DNA crystals. *Nature* **1998**, *394*, 539–544.

[104] Jungmann, R.; Scheible, M.; Simmel, F. C. Nanoscale imaging in DNA nanotechnology. *Wiley Interdiscip. Rev.: Nanomed. Nanobiotechnol.* **2012**, *4*, 66–81.

[105] Allison, D. P.; Bottomley, L. A.; Thundat, T.; Brown, G. M.; Woychik, R. P.; Schrick, J. J.; Jacobson, K. B.; Warmack, R. J. Immobilization of DNA for scanning probe microscopy. *Proc. Natl. Acad. Sci. U. S. A* **1992**, *89*, 10129–33.

[106] Shao, Z.; Mou, J.; Czajkowsky, D. M.; Yang, J.; Yuan, J.-Y. Biological atomic force microscopy: what is achieved and what is needed. *Adv. Phys.* **1996**, *45*, 1–86.

[107] Pastre, D.; Pietrement, O.; Fusil, P.; Landousy, F.; Jeusset, J.; David, M. O.; Hamon, C.; Le Cam, E.; Zozime, A. Adsorption of DNA to mica mediated by divalent counterions: A theoretical and experimental study. *Biophys. J.* **2003**, *85*, 2507–2518.

[108] `http://www.atdbio.com/tools/oligo-calculator`.

[109] Reddington, S. C.; Tippmann, E. M.; Jones, D. D. Residue choice defines efficiency and influence of bioorthogonal protein modification via genetically encoded strain promoted Click chemistry. *Chem. Commun.* **2012**, *48*, 8419–8421.

[110] Zaki, A. J.; Hartley, A. M.; Reddington, S. C.; Thomas, S. K.; Watson, P.; Hayes, A.; Moskalenko, A. V.; Craciun, M. F.; Macdonald, J. E.; Jones, D. D. Defined covalent assembly of protein molecules on graphene using a genetically encoded photochemical reaction handle. *RSC Adv.* **2018**, *8*, 5768–5775.

[111] Putman, C. A. J.; Van der Werf, K. O.; De Grooth, B. G.; Van Hulst, N. F.; Greve, J. Tapping mode atomic force microscopy in liquid. *Applied Physics Letters* **1994**, *64*, 2454–2456.

[112] Liu, Z.; Li, Z.; Zhou, H.; Wei, G.; Song, Y.; Wang, L. Imaging DNA molecules on mica surface by atomic force microscopy in air and in liquid. *Microscopy research and technique* **2005**, *66*, 179–185.

[113] Asmatulu, R.; Khan, W. S. In *Synthesis and Applications of Electrospun Nanofibers*; Asmatulu, R., Khan, W. S., Eds.; Micro and Nano Technologies; Elsevier, 2019; pp 257–281.

[114] Hansma, H.; Vesenka, J.; Siegerist, C.; Kelderman, G.; Morrett, H.; Sinsheimer, R. L.; Elings, V.; Bustamante, C.; Hansma, P. Reproducible imaging and dissection of plasmid DNA under liquid with the atomic force microscope. *Science* **1992**, *256*, 1180–1184.

[115] Yang, J.; Takeyasu, K.; Shao, Z. Atomic force microscopy of DNA molecules. *FEBS Lett.* **1992**, *301*, 173–176.

[116] Hussain, R.; Benning, K.; Javorfi, T.; Longo, E.; Rudd, T. R.; Pulford, B.; Siligardi, G. CDApps: integrated software for experimental planning and data processing at beamline B23, Diamond Light Source. *J. Synchrotron Radiat.* **2015**, *22*, 465–468.

[117] Wu, Y.; Huang, H. W.; Olah, G. A. Method of oriented circular dichroism. *Biophys. J.* **1990**, *57*, 797–806.

[118] Burck, J.; Wadhwani, P.; Fanghanel, S.; Ulrich, A. S. Oriented circular dichroism: a method to characterize membrane-active peptides in oriented lipid bilayers. *Acc. Chem. Res.* **2016**, *49*, 184–192.

[119] Baker, B. R.; Garrell, R. L. G-factor analysis of protein secondary structure in solutions and thin films. *Faraday Discuss.* **2004**, *126*, 209–222.

[120] McPhie, P. Circular dichroism studies on proteins in films and in solution: estimation of secondary structure by g-factor analysis. *Anal. Biochem.* **2001**, *293*, 109–119.

[121] Fu, T. J.; Seeman, N. C. DNA double-crossover molecules. *Biochem.* **1993**, *32*, 3211–3220.

[122] Rusling, D. A.; Nandhakumar, I. S.; Brown, T.; Fox, K. R. Triplex-directed recognition of a DNA nanostructure assembled by crossover strand exchange. *ACS nano* **2012**, *6*, 3604–3613.

[123] Liu, F.; Sha, R.; Seeman, N. C. Modifying the surface features of two-dimensional DNA crystals. *J. Am. Chem. Soc.* **1999**, *121*, 917–922.

[124] Endo, M.; Seeman, N. C.; Majima, T. DNA Tube Structures Controlled by a Four-Way-Branched DNA Connector. *Angew. Chem., Int. Ed.* **2005**, *44*, 6074–6077.

[125] Itagaki, E.; Palmer, G.; Hager, L. P. Studies on Cytochrome b562 of Escherichia coli II. RECONSTITUTION OF CYTOCHROME b562 FROM APOPROTEIN AND HEMIN. *J. Biol. Chem.* **1967**, *242*, 2272–2277.

[126] Mathews, F. S.; Bethge, P. H.; Czerwinski, E. The structure of cytochrome b562 from Escherichia coli at 2.5 A resolution. *J. Biol. Chem.* **1979**, *254*, 1699–1706.

[127] Arnesano, F.; Banci, L.; Bertini, I.; Faraone-Mennella, J.; Rosato, A.; Barker, P. D.; Fersht, A. R. The Solution Structure of Oxidized Escherichia coli Cytochrome b562. *Biochemistry* **1999**, *38*, 8657–8670.

[128] Feng, Y.; Sligar, S. G.; Wand, A. J. Solution structure of apocytochrome b 562. *Nat. Struct. Biol.* **1994**, *1*, 30–35.

[129] Low, D. W.; Hill, M. G.; Carrasco, M. R.; Kent, S. B.; Botti, P. Total synthesis of cytochrome b562 by native chemical ligation using a removable auxiliary. *Proc. Natl. Acad. Sci.* **2001**, *98*, 6554–6559.

[130] Bullock, P. A.; Myer, Y. P. Circular dichroism and resonance Raman studies of cytochrome b562 from Escherichia coli. *Biochemistry* **1978**, *17*, 3084–3091.

[131] Robinson, C. R.; Sligar, S. G. Electrostatic stabilization in four-helix bundle proteins. *Protein Sci.* **1993**, *2*, 826–837.

[132] Kamiya, N.; Okimoto, Y.; Ding, Z.; Ohtomo, H.; Shimizu, M.; Kitayama, A.; Morii, H.; Nagamune, T. How does heme axial ligand deletion affect the structure and the function of cytochrome b 562? *Protein Eng.* **2001**, *14*, 415–419.

[133] Cary, P.; Kneale, G. G. *DNA-Protein Interactions*; Springer, 2009; pp 613–624.

[134] Johnson, B. B.; Dahl, K. S.; Tinoco Jr, I.; Ivanov, V. I.; Zhurkin, V. B. Correlations between deoxyribonucleic acid structural parameters and calculated circular dichroism spectra. *Biochem.* **1981**, *20*, 73–78.

[135] Bumgarner, R. Overview of DNA microarrays: types, applications, and their future. *Curr. Protoc. Mol. Biol.* **2013**, *101*, 22.1. 1–22.1. 11.

[136] Nimse, S. B.; Song, K.; Sonawane, M. D.; Sayyed, D. R.; Kim, T. Immobilization techniques for microarray: challenges and applications. *Sensors* **2014**, *14*, 22208–22229.

[137] Ko, K.-S.; Jaipuri, F. A.; Pohl, N. L. Fluorous-based carbohydrate microarrays. *J. Am. Chem. Soc.* **2005**, *127*, 13162–13163.

[138] Flynn, G. E.; Withers, J. M.; Macias, G.; Sperling, J. R.; Henry, S. L.; Cooper, J. M.; Burley, G. A.; Clark, A. W. Reversible DNA micro-patterning using the fluorous effect. *Chem. Commun.* **2017**, *53*, 3094–3097.

[139] Nicholson, R. L.; Ladlow, M. L.; Spring, D. R. Fluorous tagged small molecule microarrays. *Chem. Commun.* **2007**, 3906–3908.

[140] Gladysz, J. A.; Curran, D. P. Fluorous chemistry: from biphasic catalysis to a parallel chemical universe and beyond. *Tetrahedron* **2002**, *58*, 3823–3825.

[141] Horváth, I. T.; Rábai, J. Facile catalyst separation without water: fluorous biphase hydroformylation of olefins. *Science* **1994**, *266*, 72–75.

[142] Gladysz, J. A.; Curran, D. P.; Horváth, I. T. *Handbook of fluorous chemistry*; John Wiley & Sons, 2006.

[143] Bégué, J.-P.; Bonnet-Delpon, D. *Bioorganic and medicinal chemistry of fluorine*; John Wiley & Sons, 2008.

[144] Krafft, M. P. Fluorocarbons and fluorinated amphiphiles in drug delivery and biomedical research. *Adv. Drug Delivery Rev.* **2001**, *47*, 209–228.

[145] Pauling, L., et al. *The Nature of the Chemical Bond...*; Cornell university press Ithaca, NY, 1960; Vol. 260.

[146] Gladysz, J. A.; Jurisch, M. *Fluorous Chem.*; Springer, 2011; pp 1–23.

[147] Studer, A.; Hadida, S.; Ferritto, R.; Kim, S.-Y.; Jeger, P.; Wipf, P.; Curran, D. P. Fluorous synthesis: A fluorous-phase strategy for improving separation efficiency in organic synthesis. *Science* **1997**, *275*, 823–826.

[148] Curran, D. P.; Hadida, S.; He, M. Thermal allylations of aldehydes with a fluorous allylstannane. Separation of organic and fluorous products by solid phase extraction with fluorous reverse phase silica gel. *The Journal of Organic Chemistry* **1997**, *62*, 6714–6715.

[149] Brittain, S. M.; Ficarro, S. B.; Brock, A.; Peters, E. C. Enrichment and analysis of peptide subsets using fluorous affinity tags and mass spectrometry. *Nat. Biotechnol.* **2005**, *23*, 463–468.

[150] Luo, Z.; Zhang, Q.; Oderaotoshi, Y.; Curran, D. P. Fluorous mixture synthesis: a fluorous-tagging strategy for the synthesis and separation of mixtures of organic compounds. *Science* **2001**, *291*, 1766–1769.

[151] Horváth, I. T. *Fluorous chemistry*; Springer Science & Business Media, 2011; Vol. 308.

[152] Kobrin, B.; Zhang, T.; Grimes, M.; Chong, K.; Wanebo, M.; Chinn, J.; Nowak, R. An improved chemical resistance and mechanical durability of hydrophobic FDTS coatings. J. Phys.: Conf. Ser. pp 454–457.

[153] Flynn, G. *Immobilisation of DNA Using the Fluorous Effect*; University of Glasgow, 2018.

[154] Matsuoka, Y.; Norden, B. Linear dichroism studies of nucleic acid bases in stretched poly (vinyl alcohol) film. Molecular orientation and electronic transition moment directions. *J. Phys. Chem.* **1982**, *86*, 1378–1386.

[155] Tinoco Jr, I.; Turner, D. H. Fluorescence detected circular dichroism. Theory. *Am. Chem. J.* **1976**, *98*, 6453–6456.

[156] Nehira, T.; Ishihara, K.; Matsuo, K.; Izumi, S.; Yamazaki, T.; Ishida, A. A sensitive method based on fluorescence-detected circular dichroism for protein local structure analysis. *Anal. Biochem.* **2012**, *430*, 179–184.

[157] Chen, G. C.; Yang, J. T. Two-point calibration of circular dichrometer with d-10-camphorsulfonic acid. *Anal. Lett.* **1977**, *10*, 1195–1207.

Appendices

Appendix A

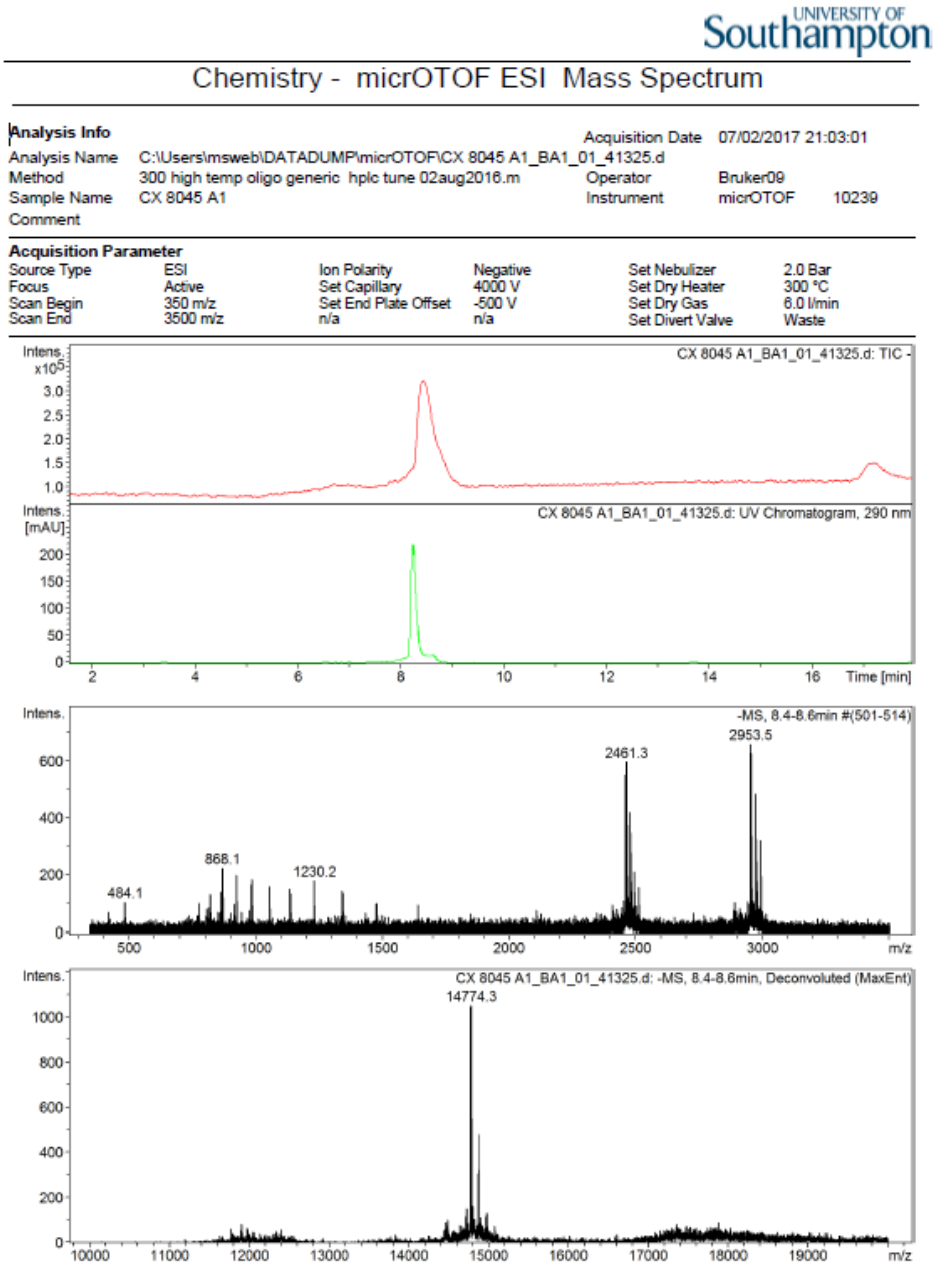

Figure 1: Mass Spectrometry analysis for CT-A1 ssDNA.

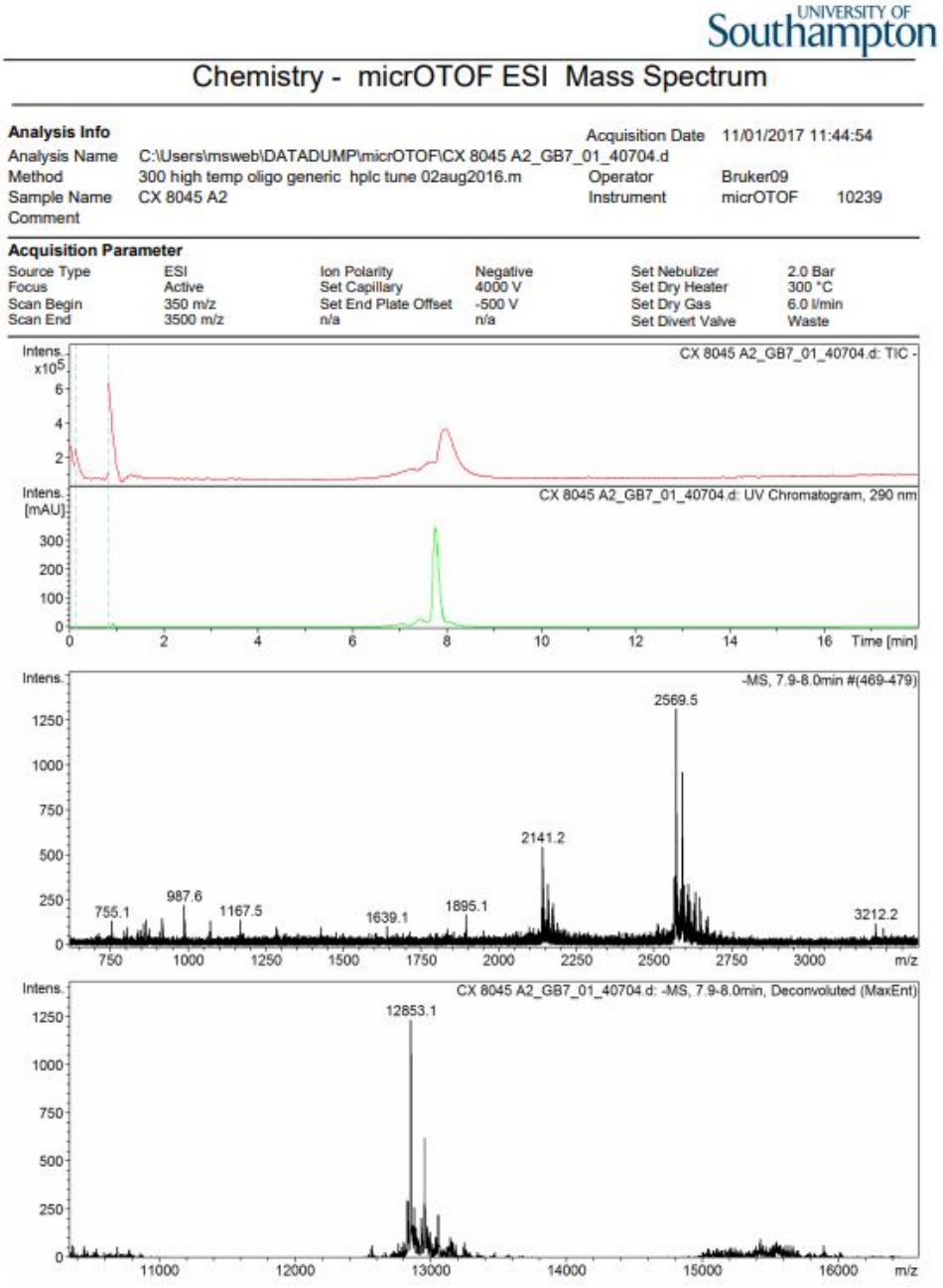

Figure 2: Mass Spectrometry analysis for CT-A2 ssDNA.

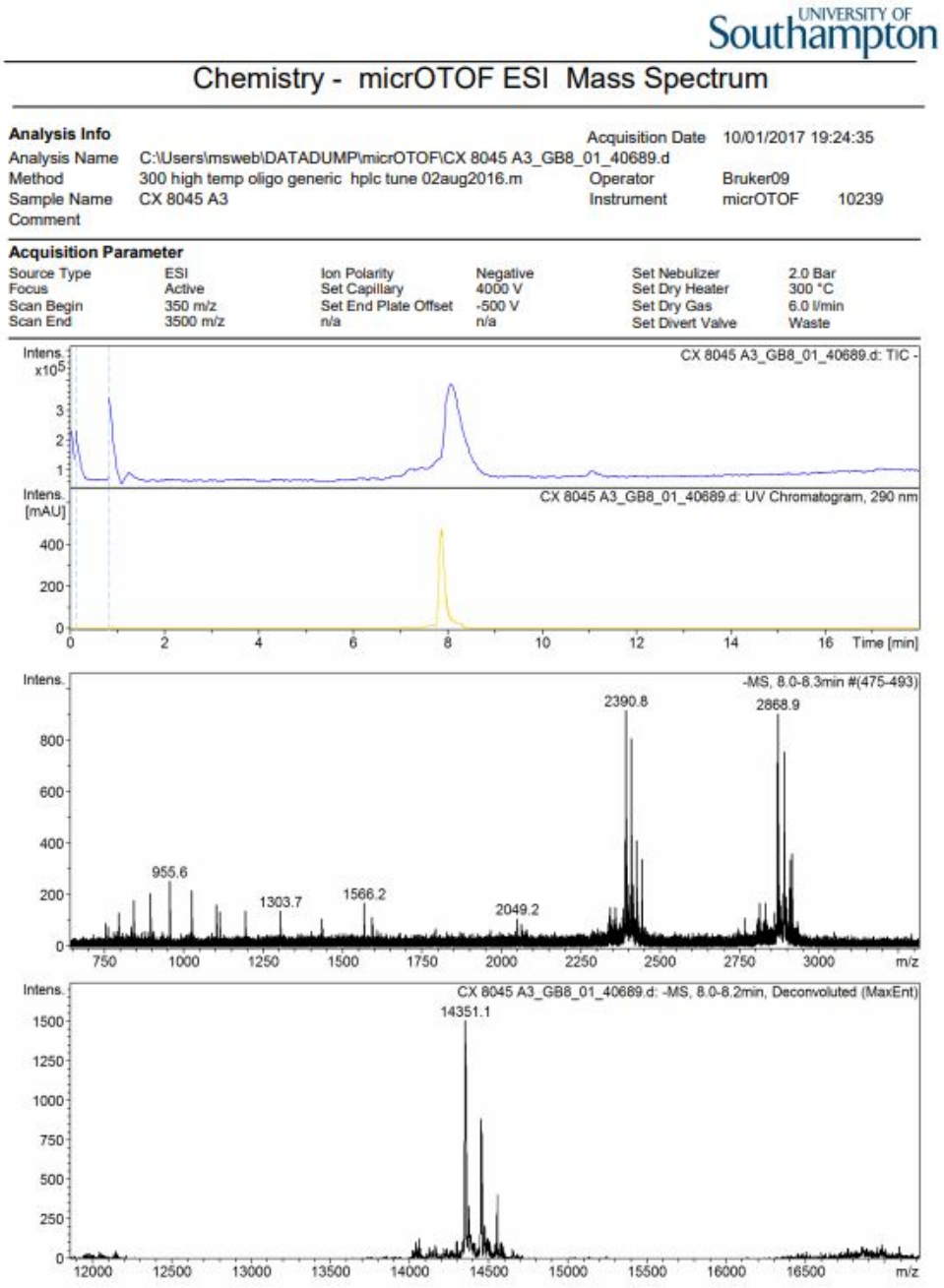

Figure 3: Mass Spectrometry analysis for CT-A3 ssDNA.

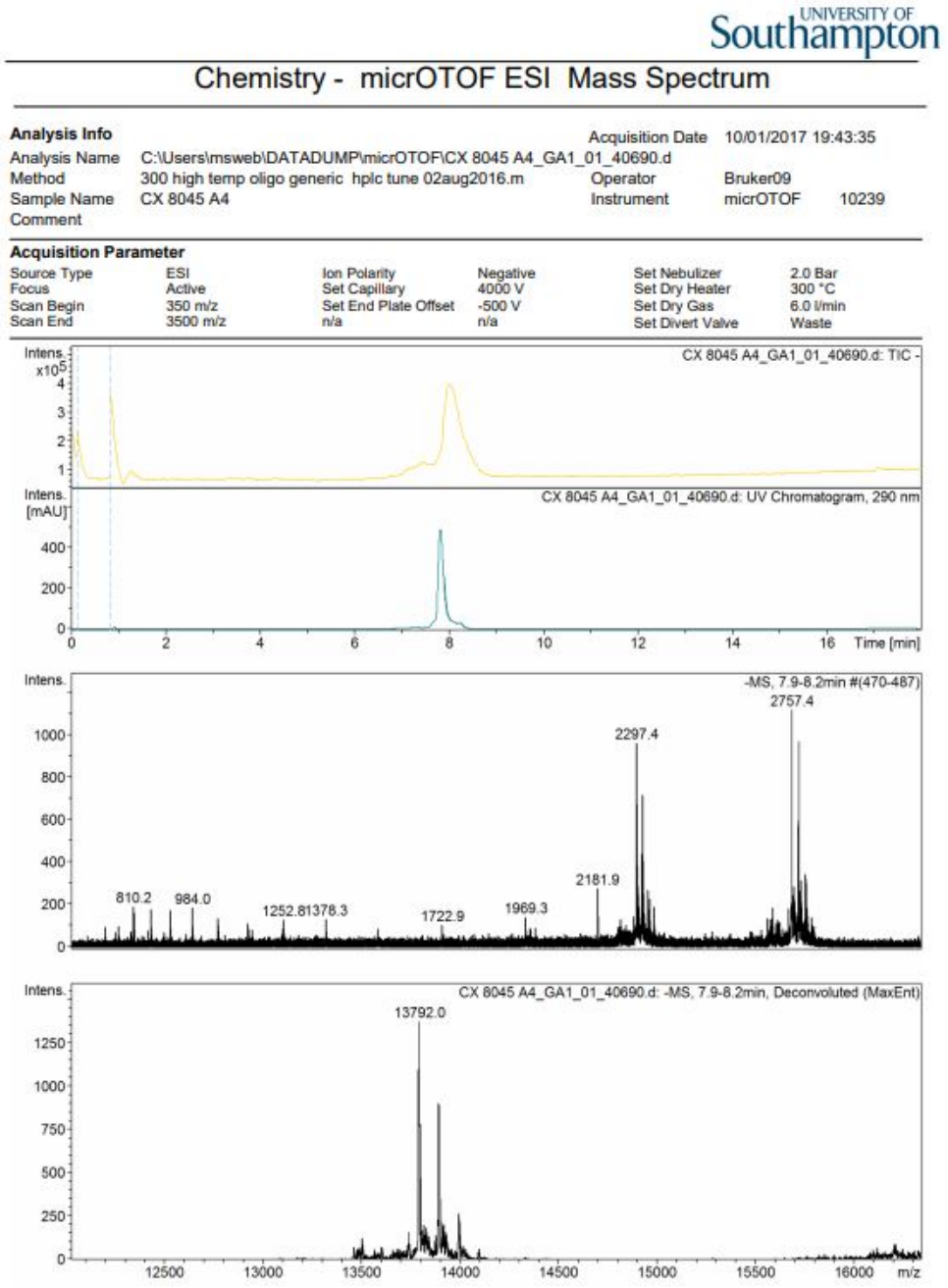

Figure 4: Mass Spectrometry analysis for CT-A4 ssDNA.

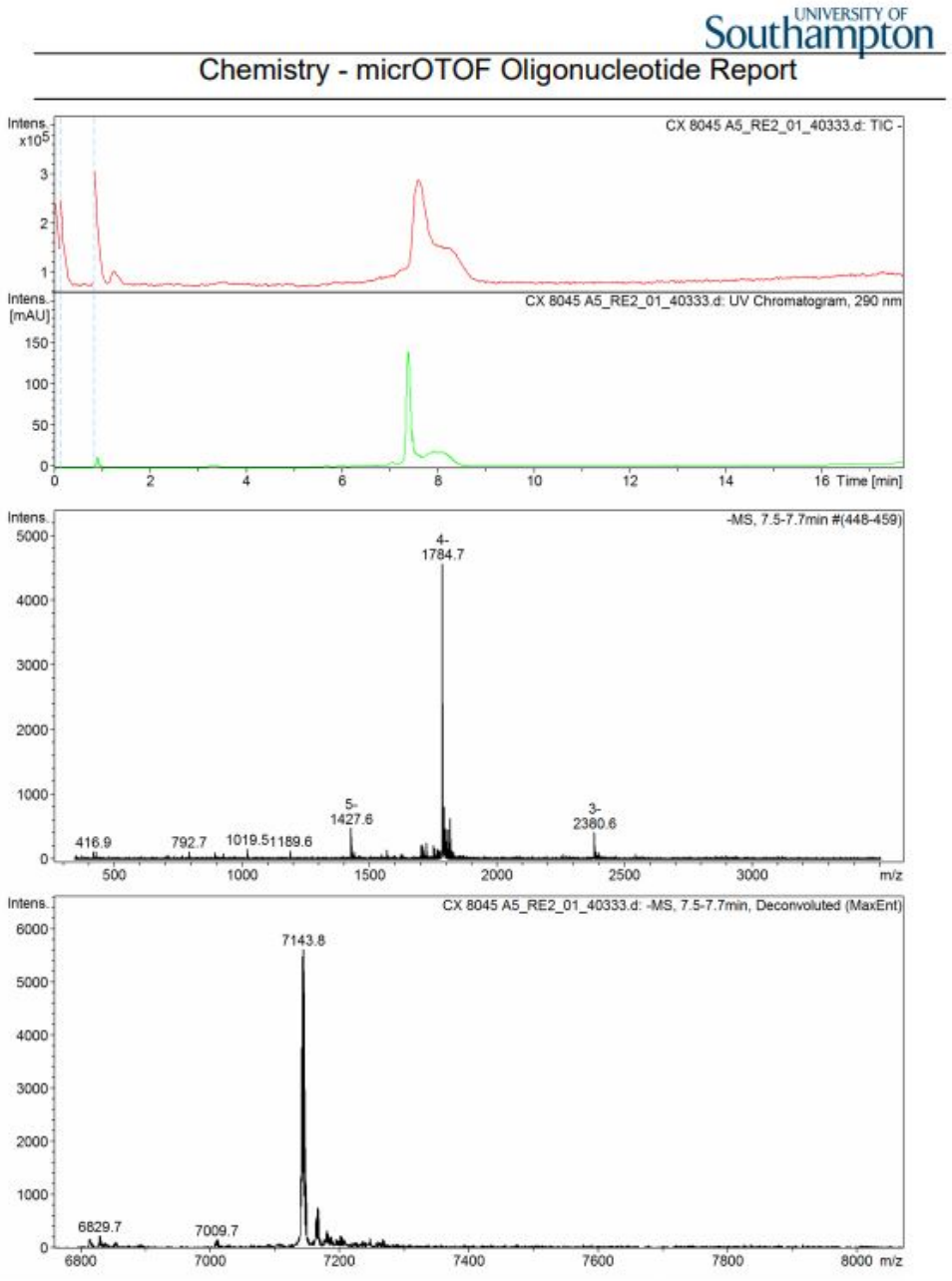

Figure 5: Mass Spectrometry analysis for CT-A5 ssDNA.

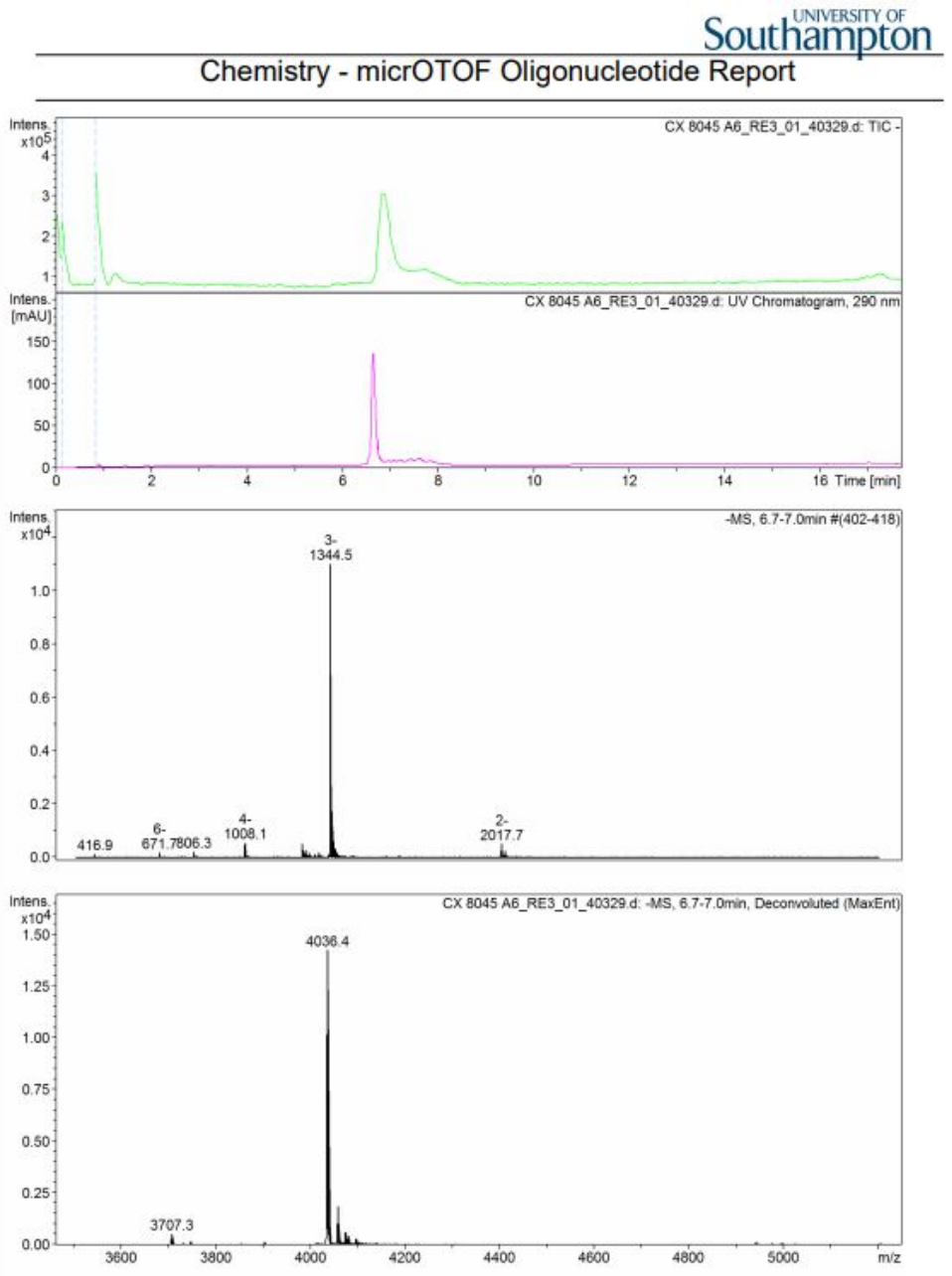

Figure 6: Mass Spectrometry analysis for CT-A6 ssDNA.

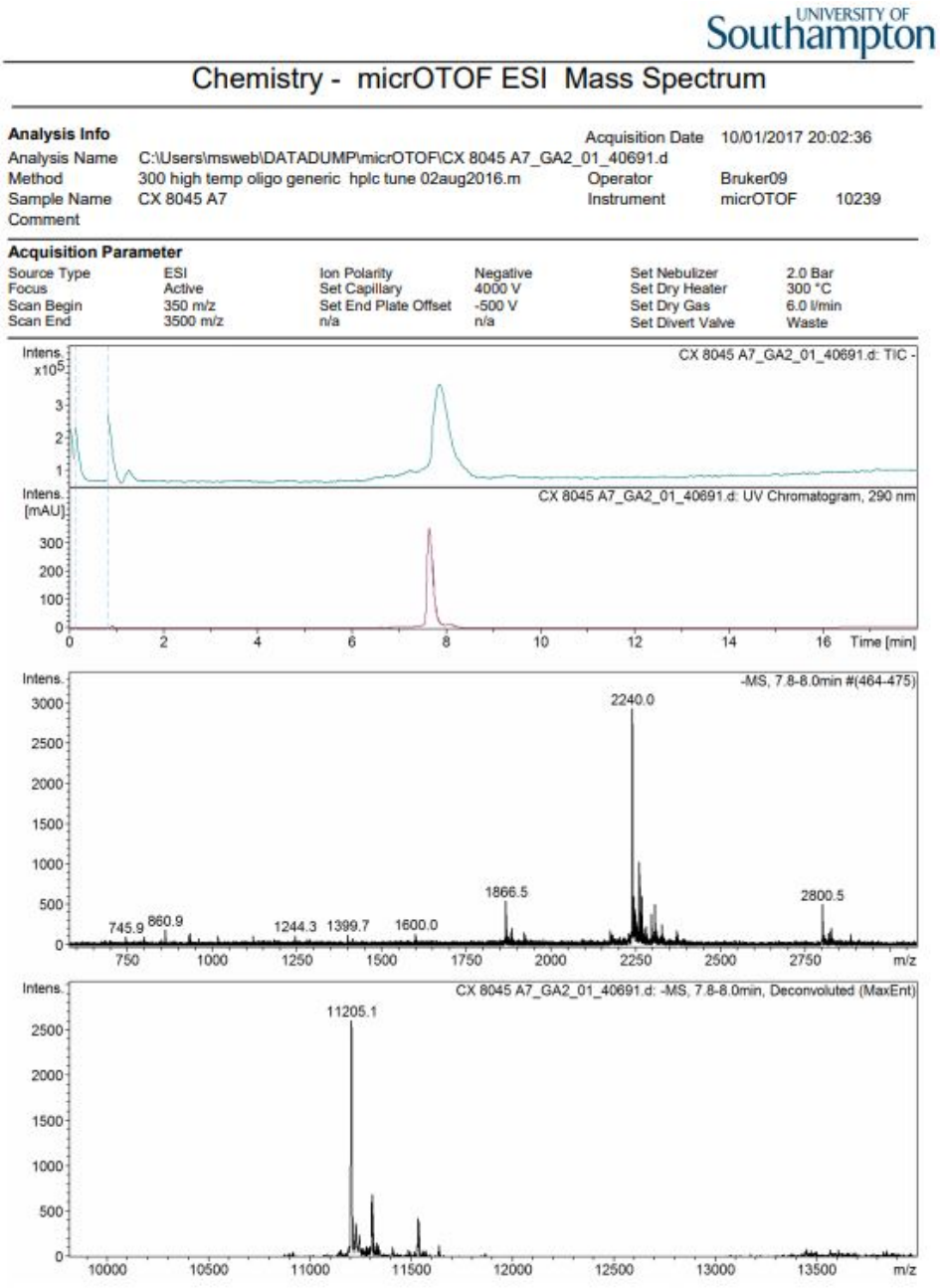

Figure 7: Mass Spectrometry analysis for CT-A7 ssDNA.

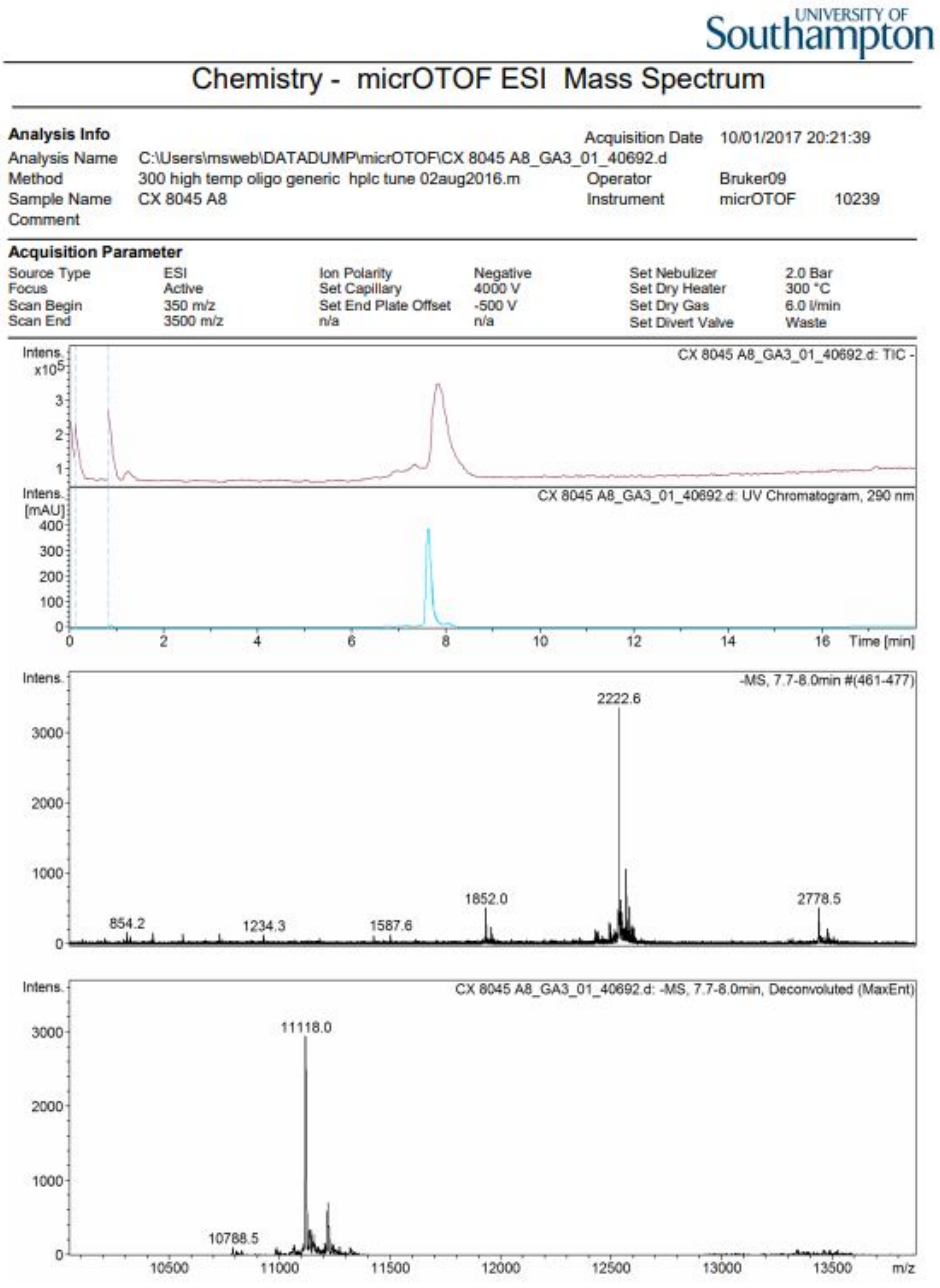

Figure 8: Mass Spectrometry analysis for CT-A8 ssDNA.

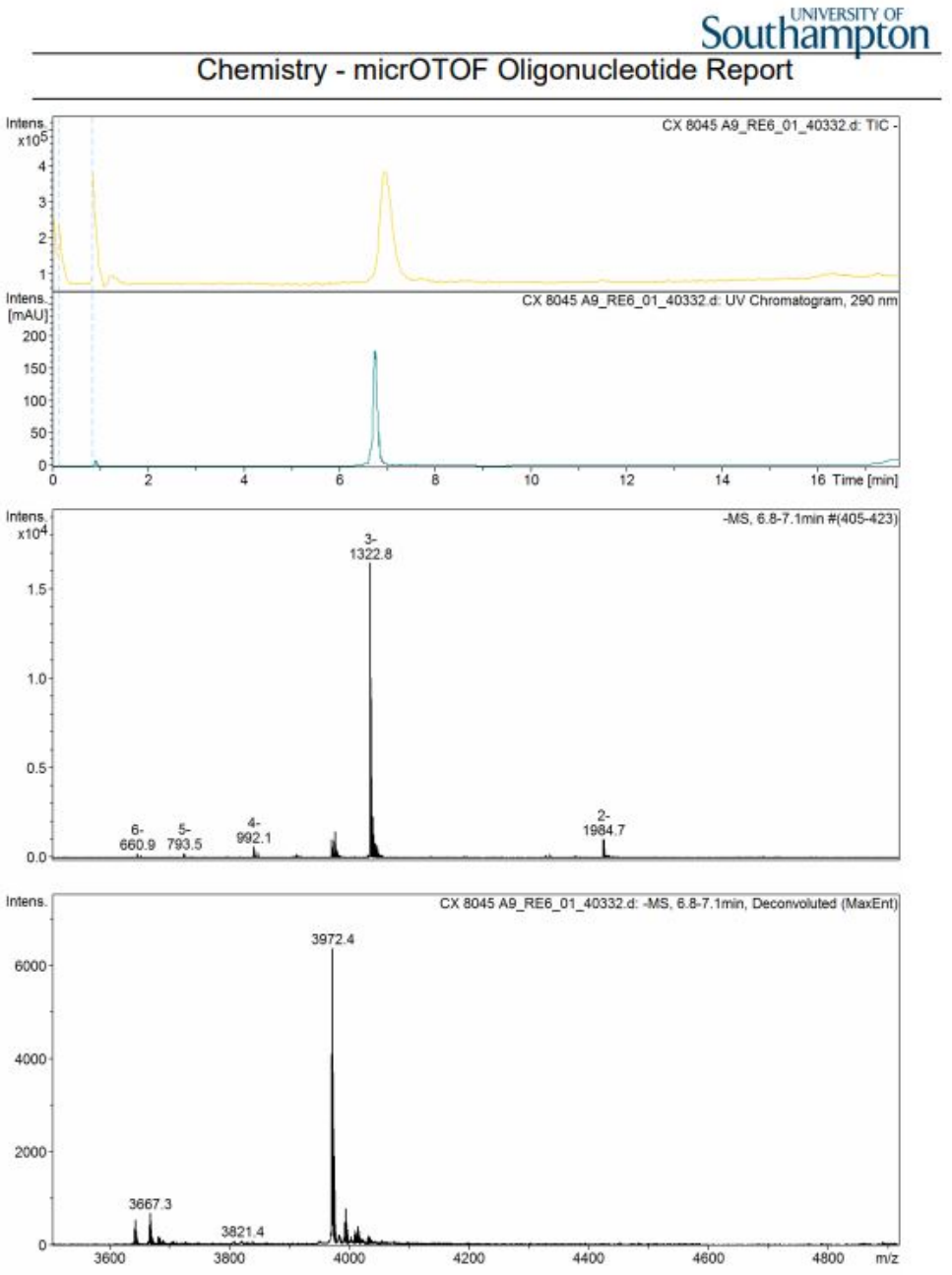

Figure 9: Mass Spectrometry analysis for CT-A9 ssDNA.

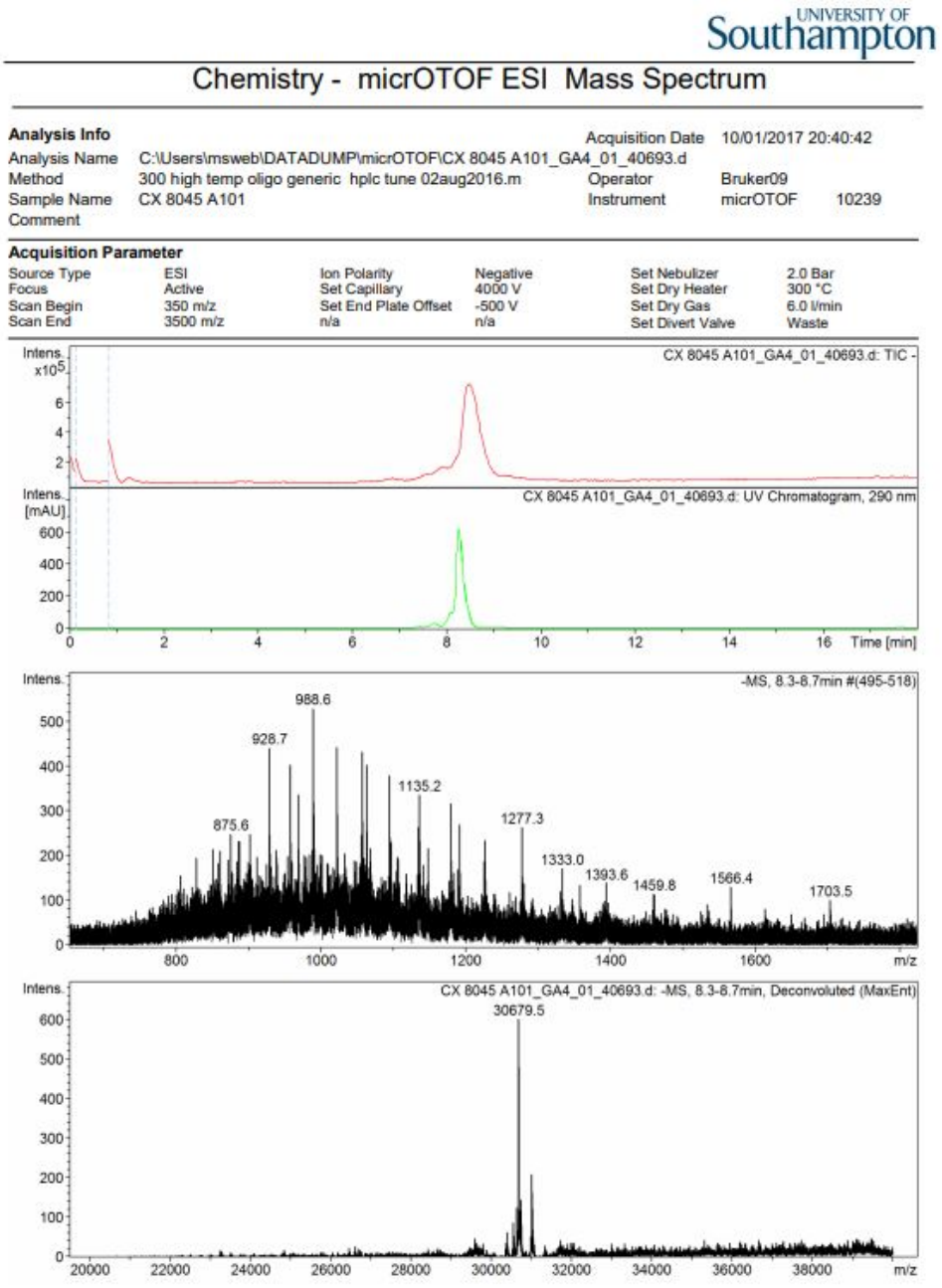

Figure 10: Mass Spectrometry analysis for CT-A10 ssDNA.

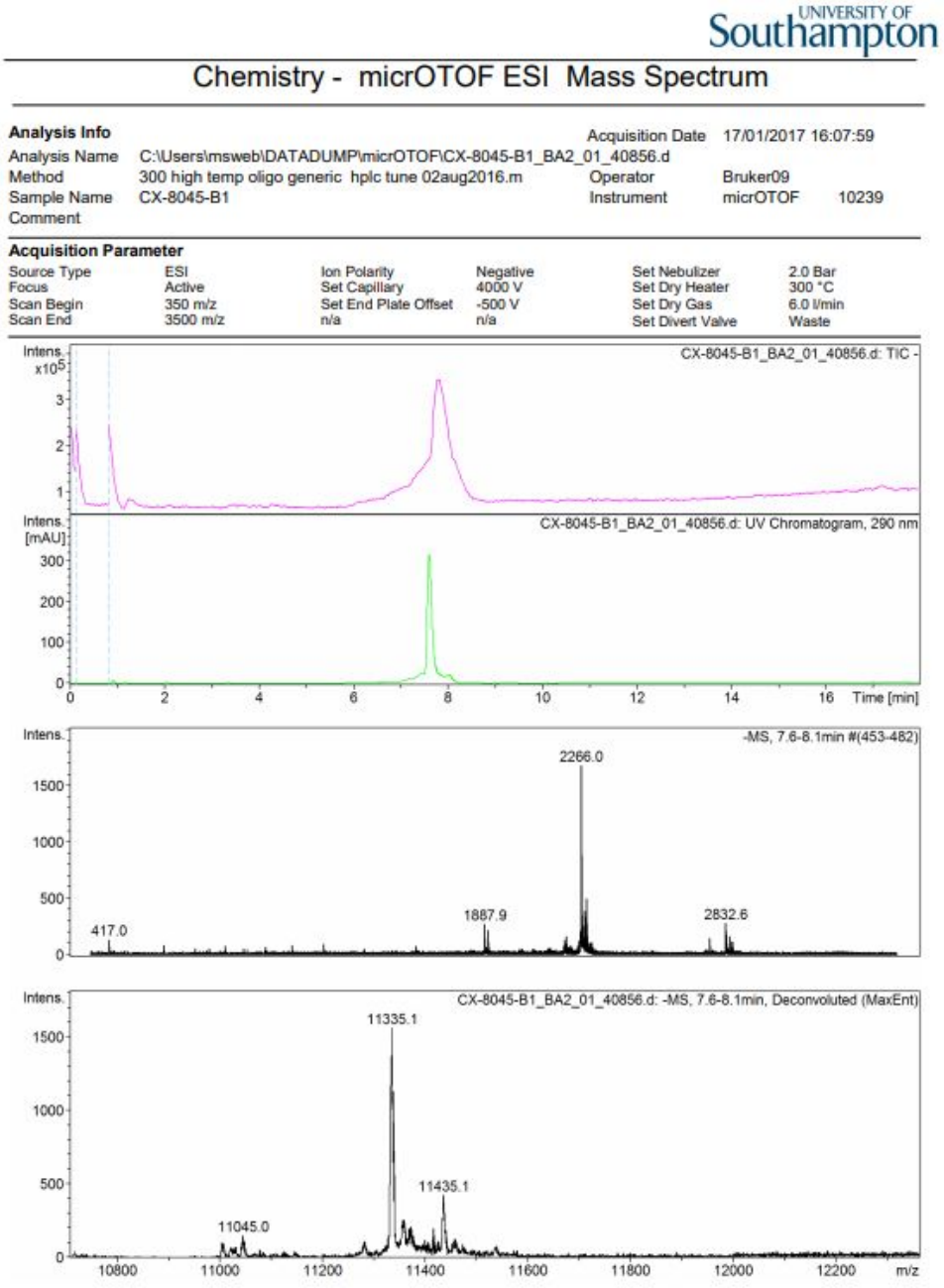

Figure 11: Mass Spectrometry analysis for CT-B1 ssDNA.

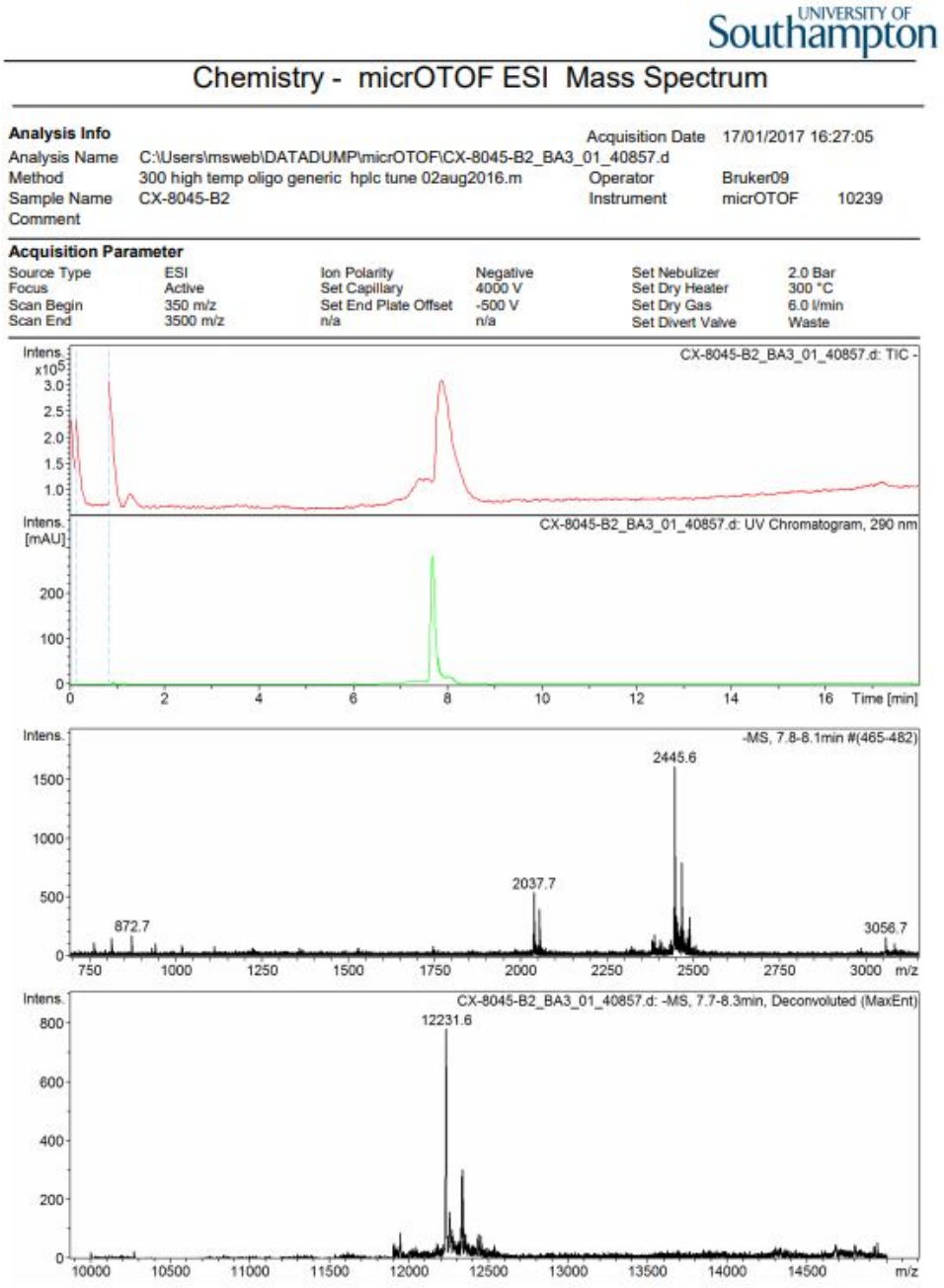

Figure 12: Mass Spectrometry analysis for CT-B2 ssDNA.

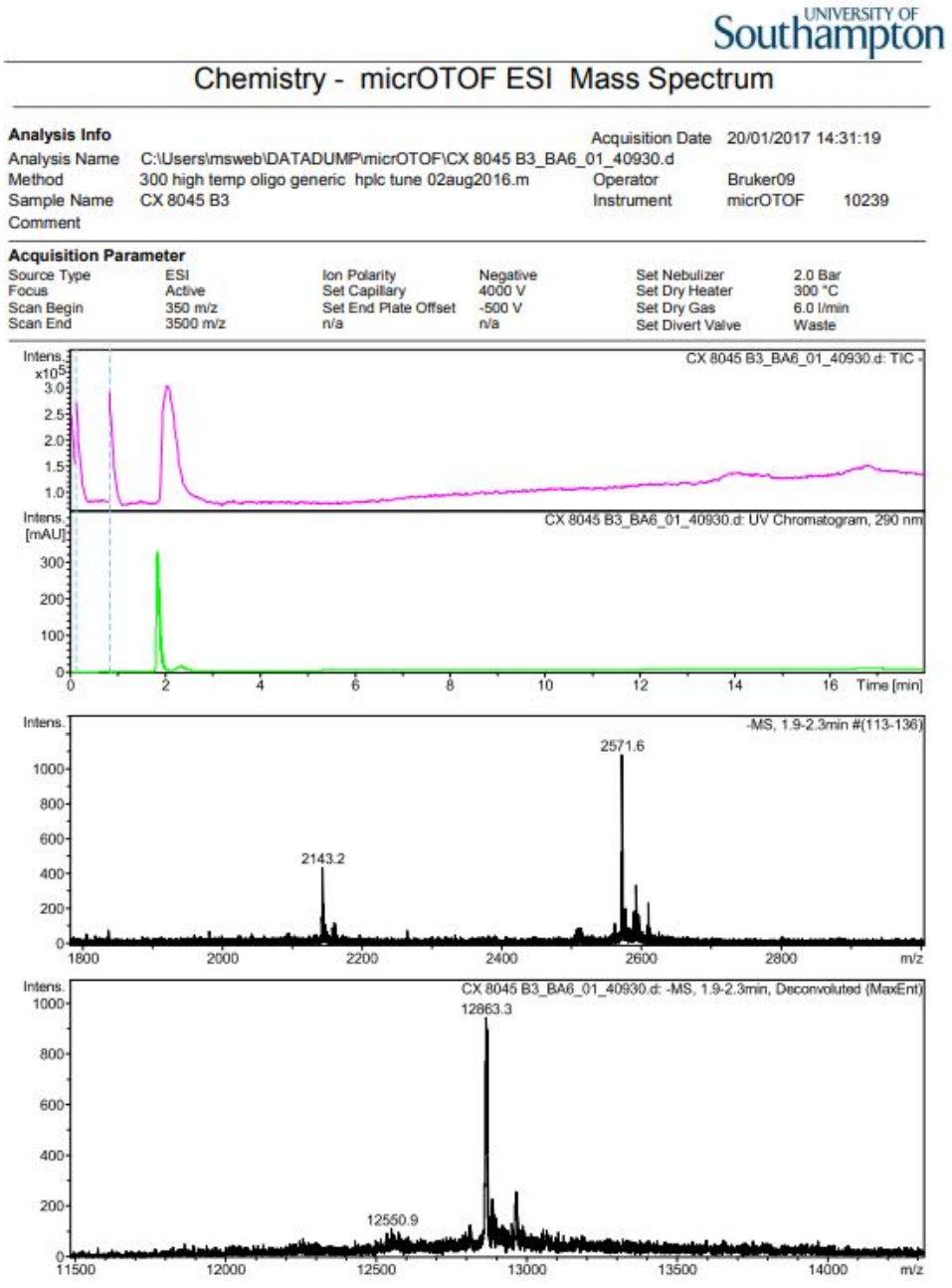

Figure 13: Mass Spectrometry analysis for CT-B3 ssDNA.

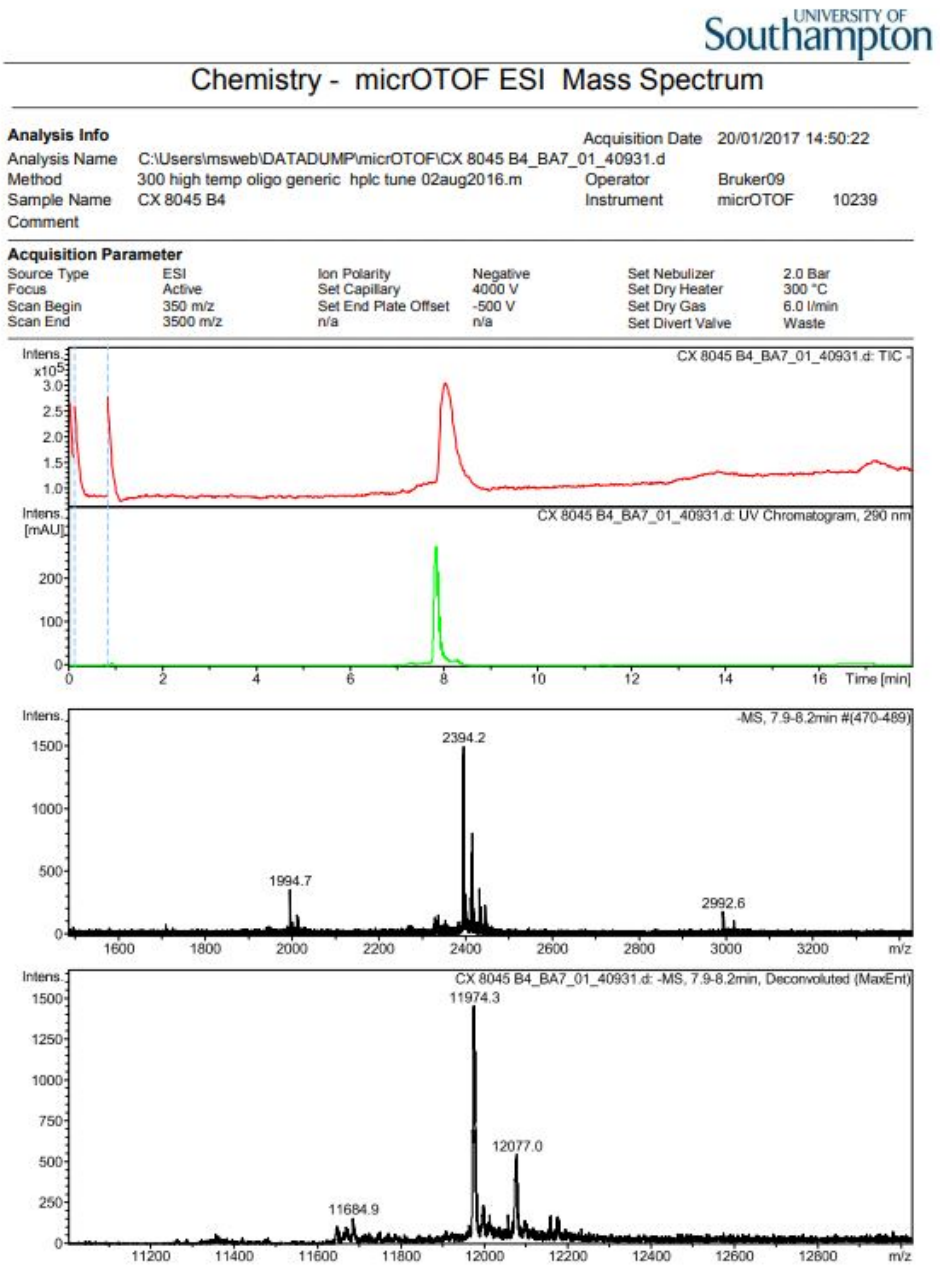

Figure 14: Mass Spectrometry analysis for CT-B4 ssDNA.

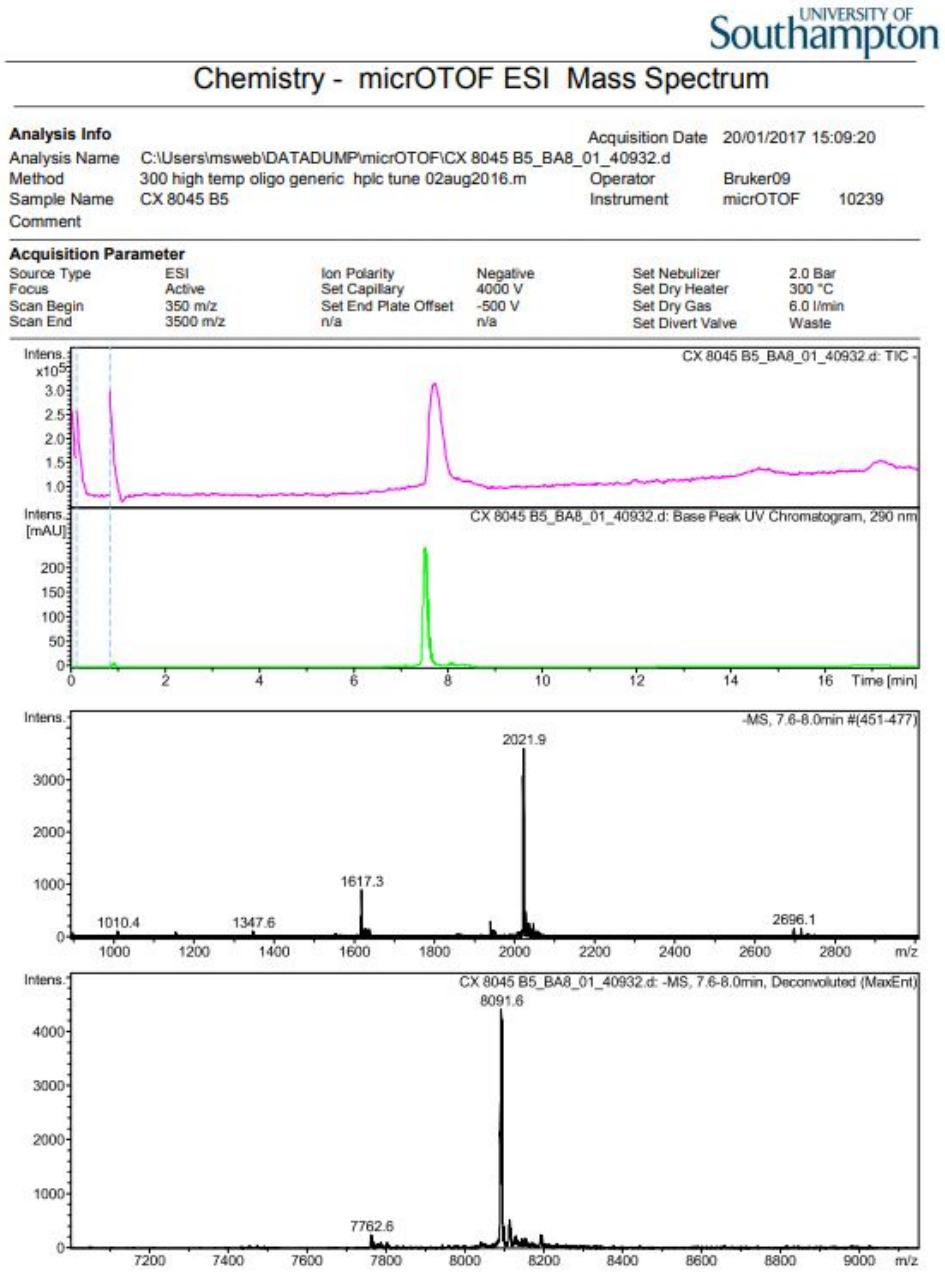

Figure 15: Mass Spectrometry analysis for CT-B5 ssDNA.

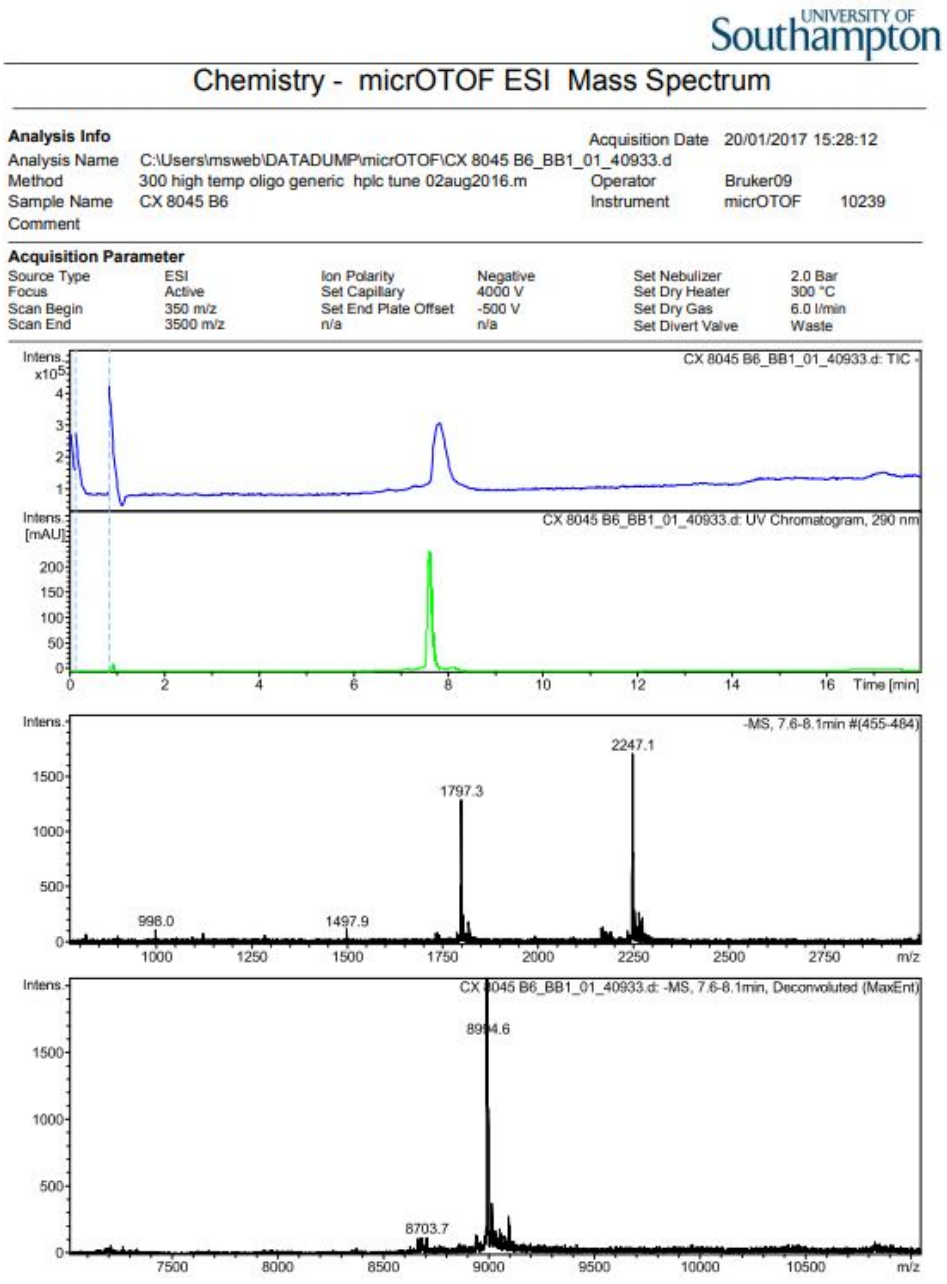

Figure 16: Mass Spectrometry analysis for CT-B6 ssDNA.

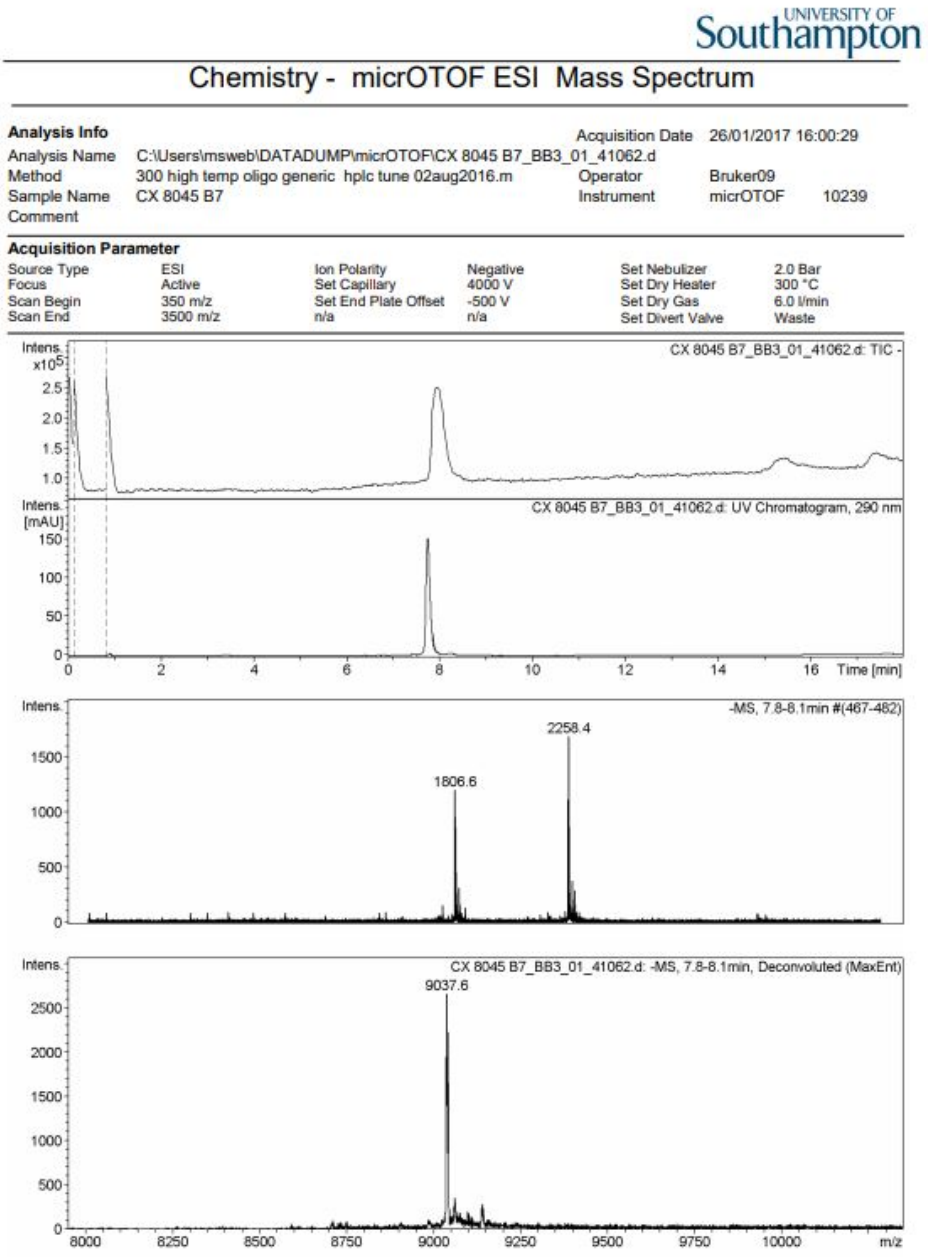

Figure 17: Mass Spectrometry analysis for CT-B7 ssDNA.

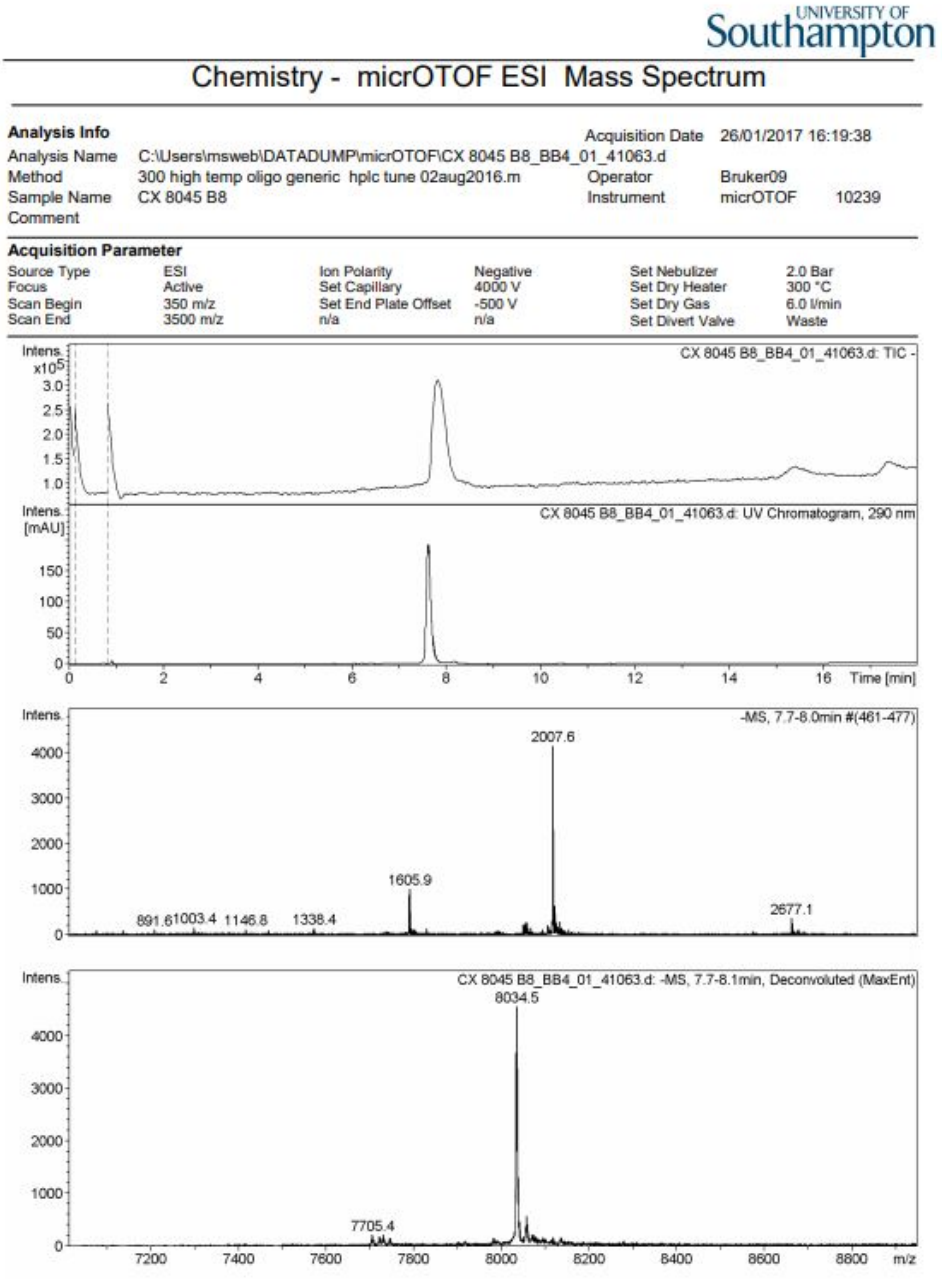

Figure 18: Mass Spectrometry analysis for CT-B8 ssDNA.

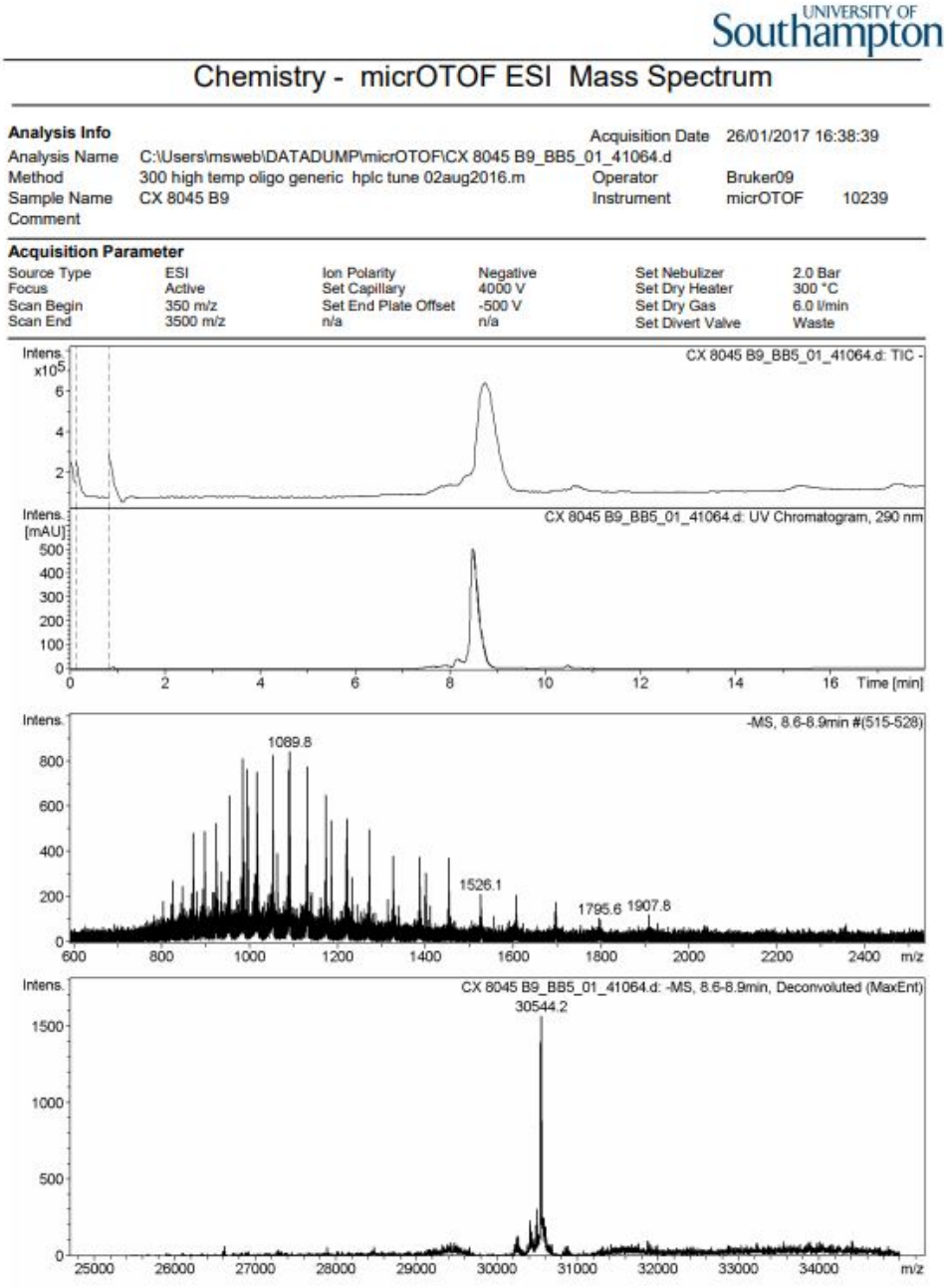

Figure 19: Mass Spectrometry analysis for CT-B9 ssDNA.

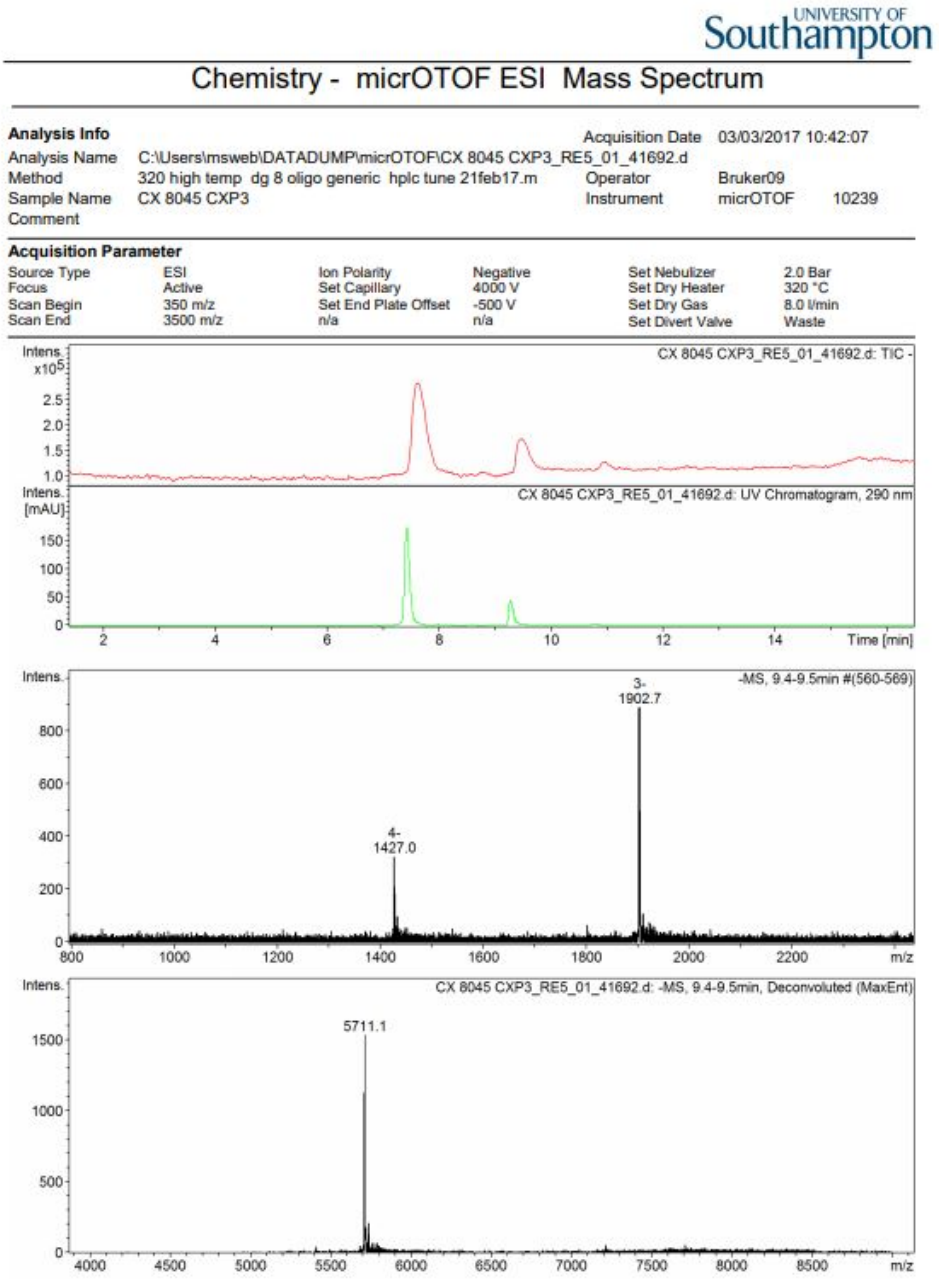

Figure 20: Mass Spectrometry analysis for the BCN-DNA.

Appendix B

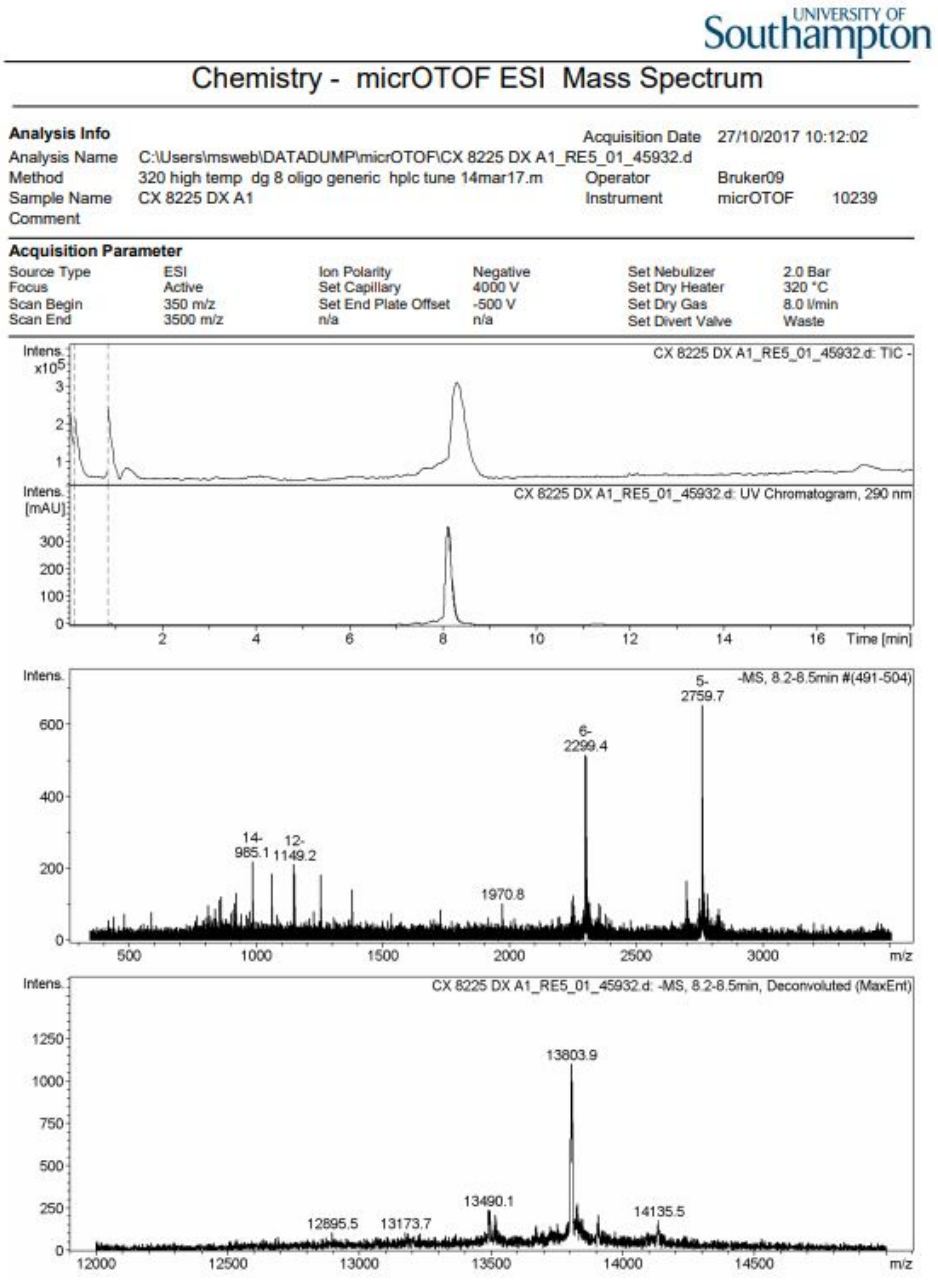

Figure 21: Mass Spectrometry analysis for DX-A1 ssDNA.

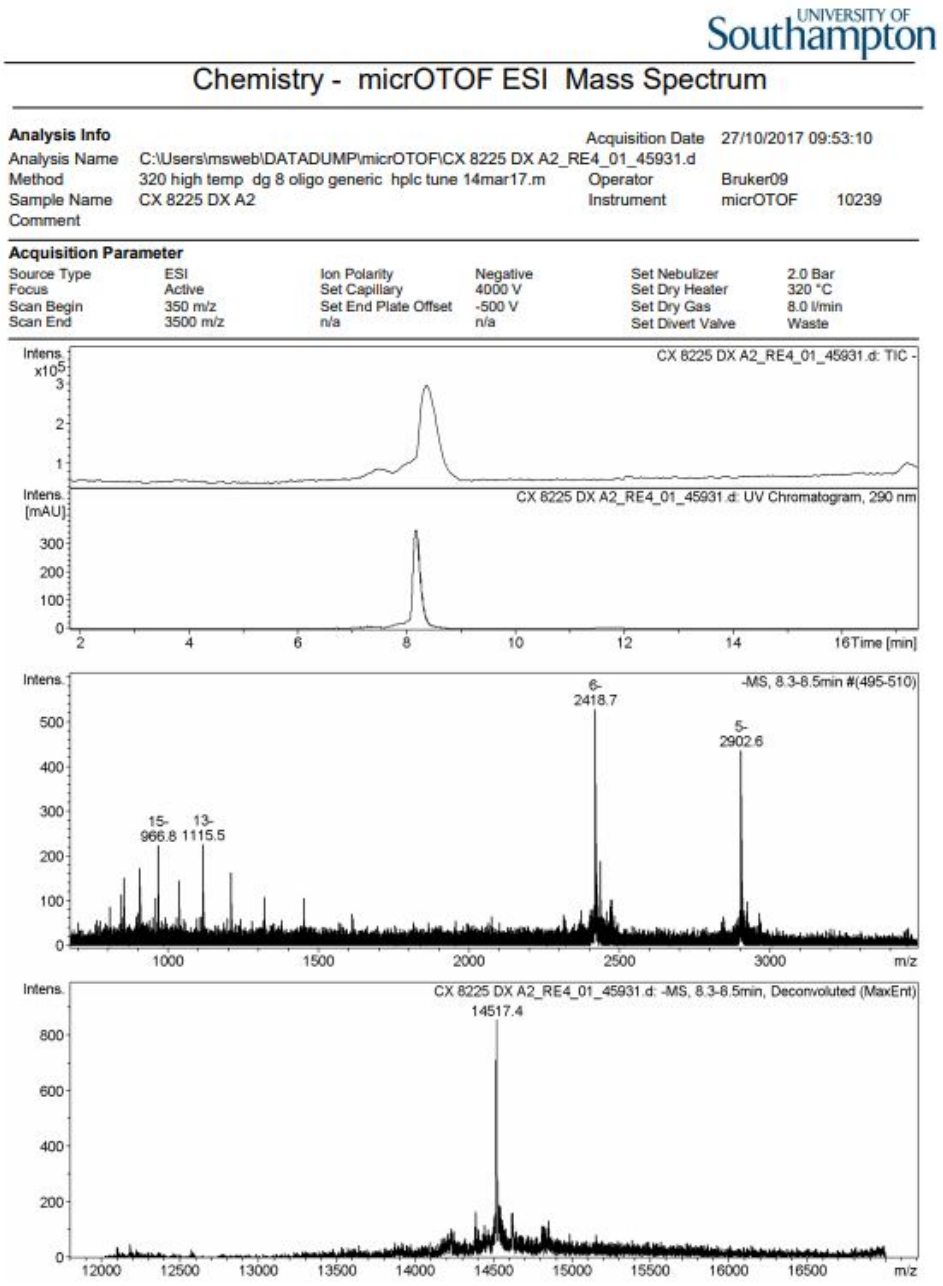

Figure 22: Mass Spectrometry analysis for DX-A2 ssDNA.

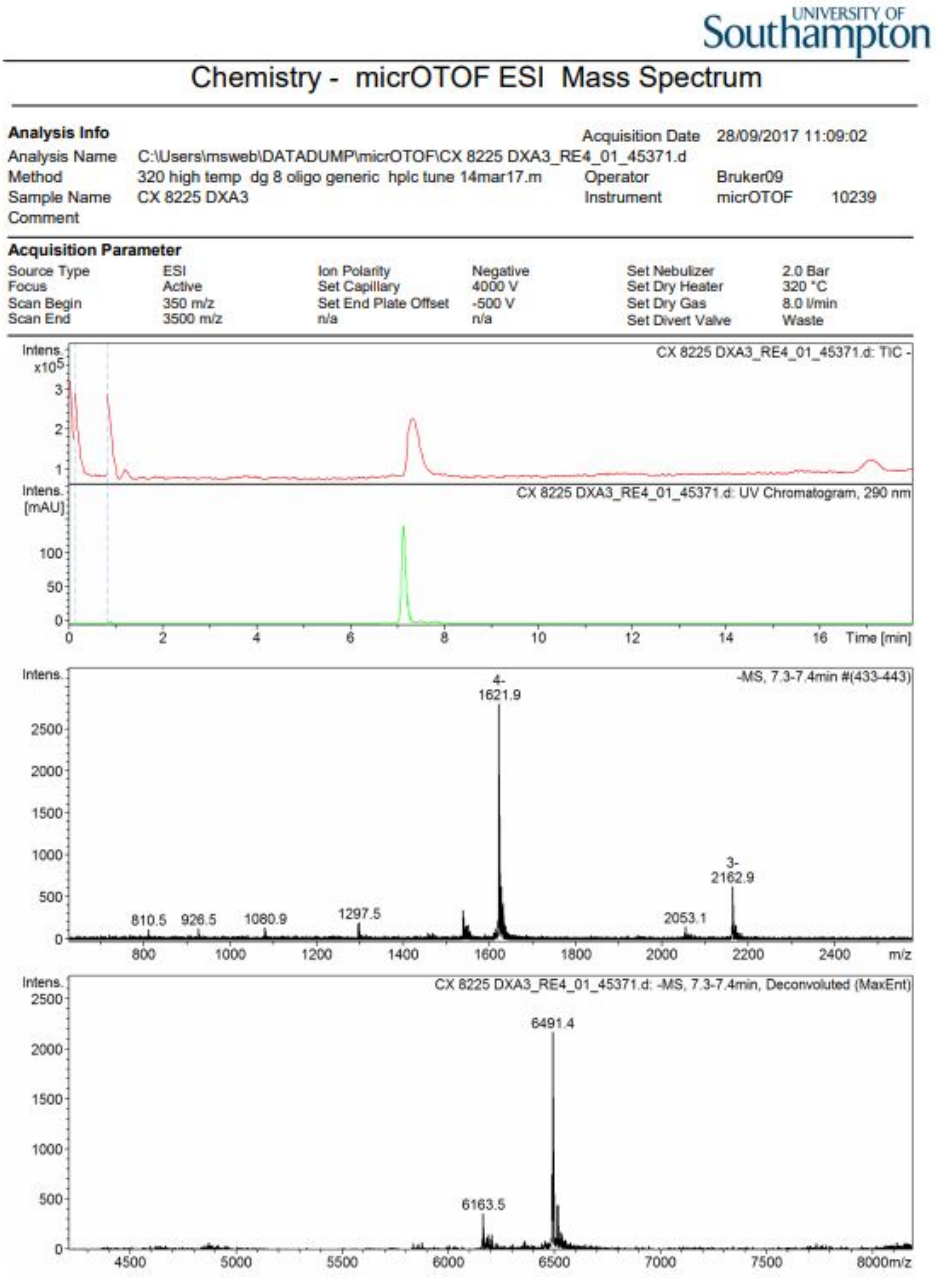

Figure 23: Mass Spectrometry analysis for DX-A3 ssDNA.

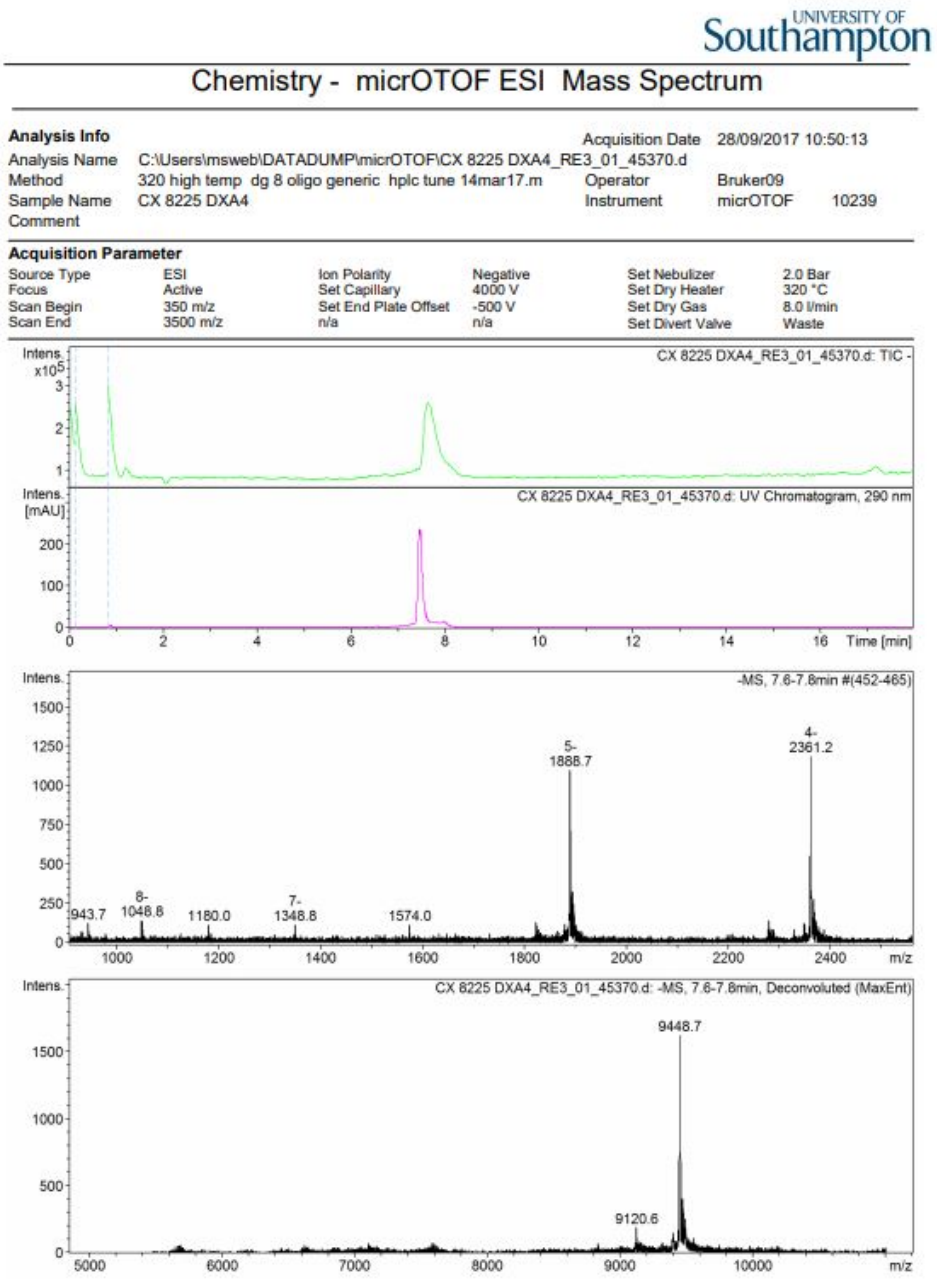

Figure 24: Mass Spectrometry analysis for DX-A4 ssDNA.

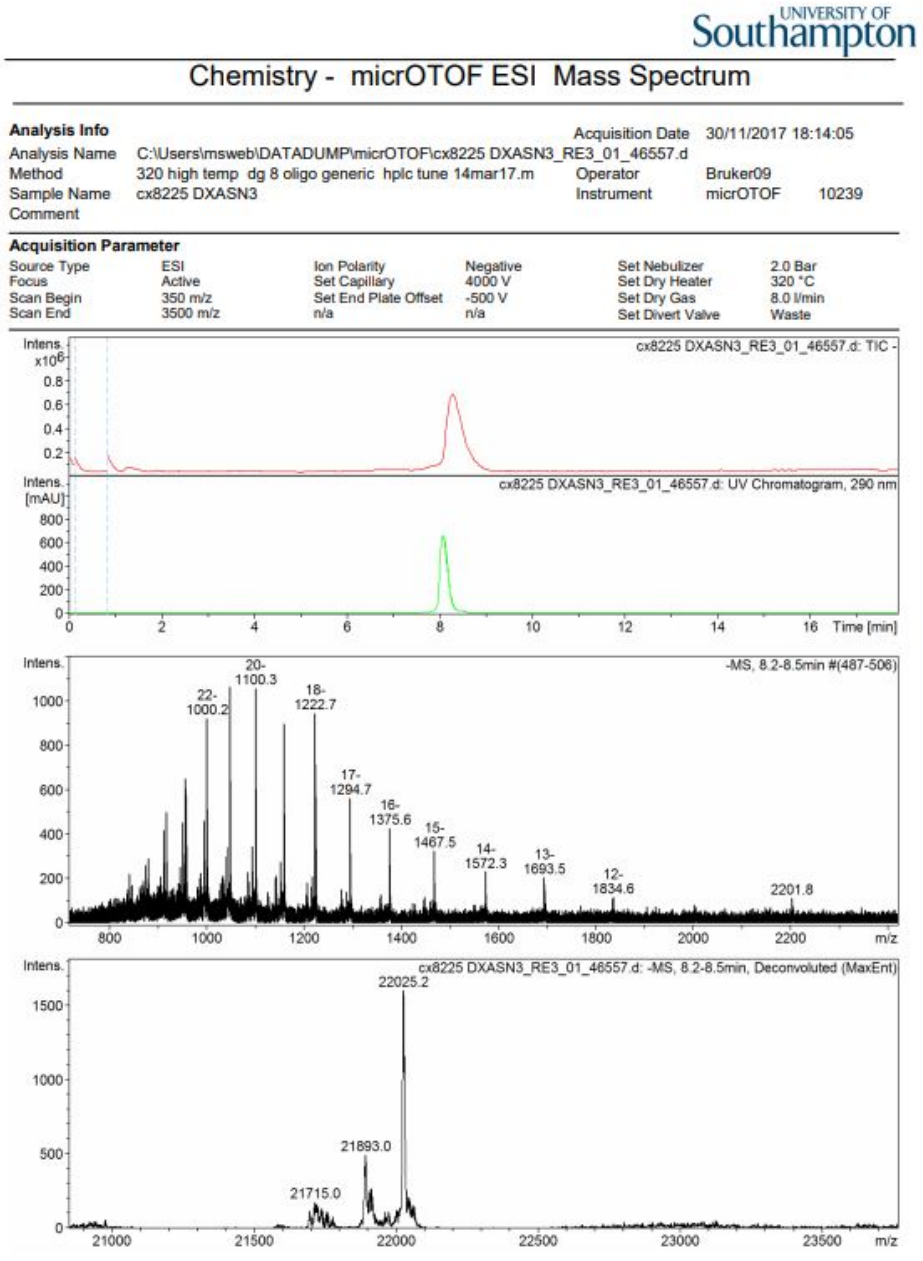

Figure 25: Mass Spectrometry analysis for DX-A5_5, ssDNA.

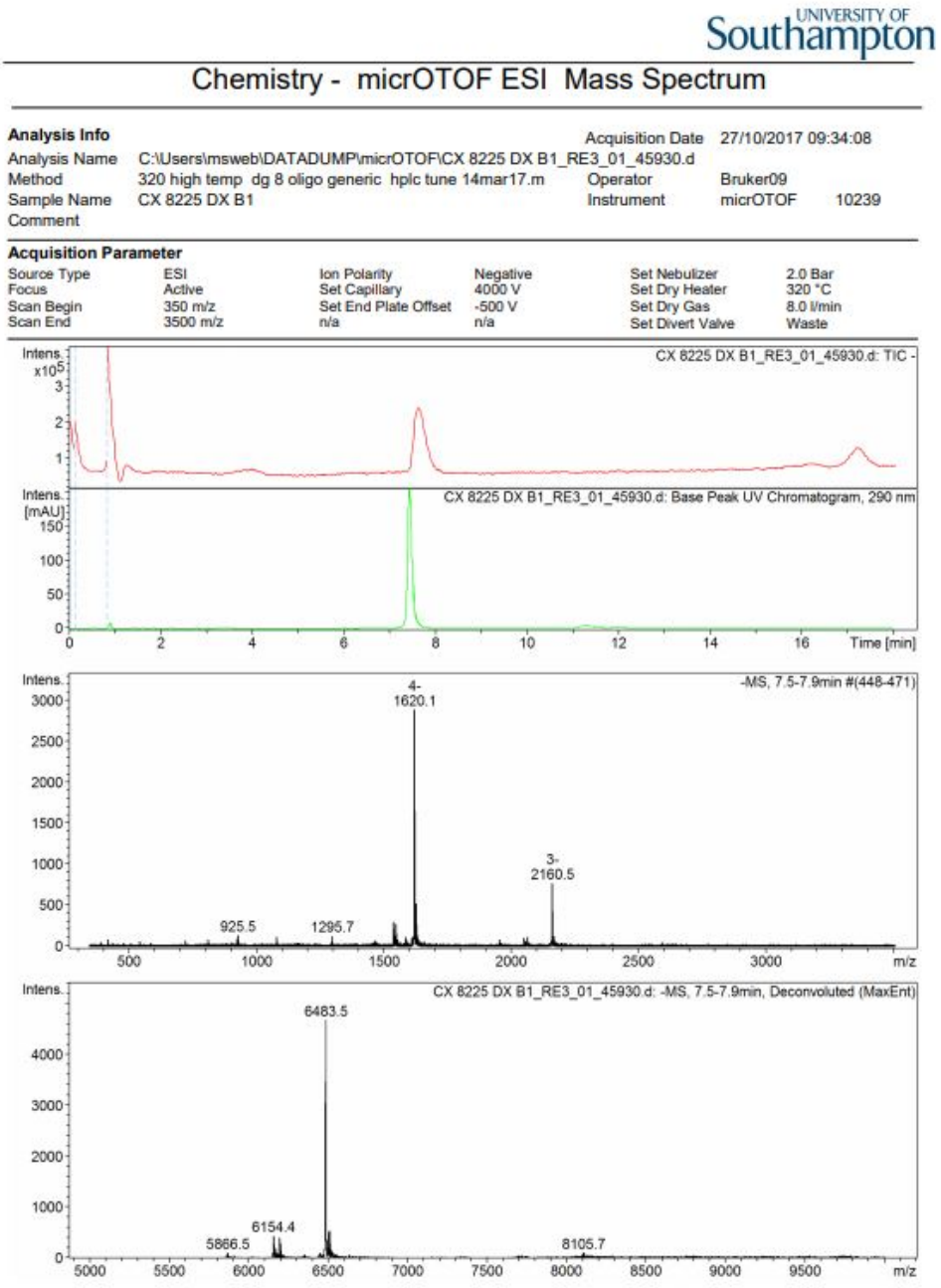

Figure 26: Mass Spectrometry analysis for DX-B1 ssDNA.

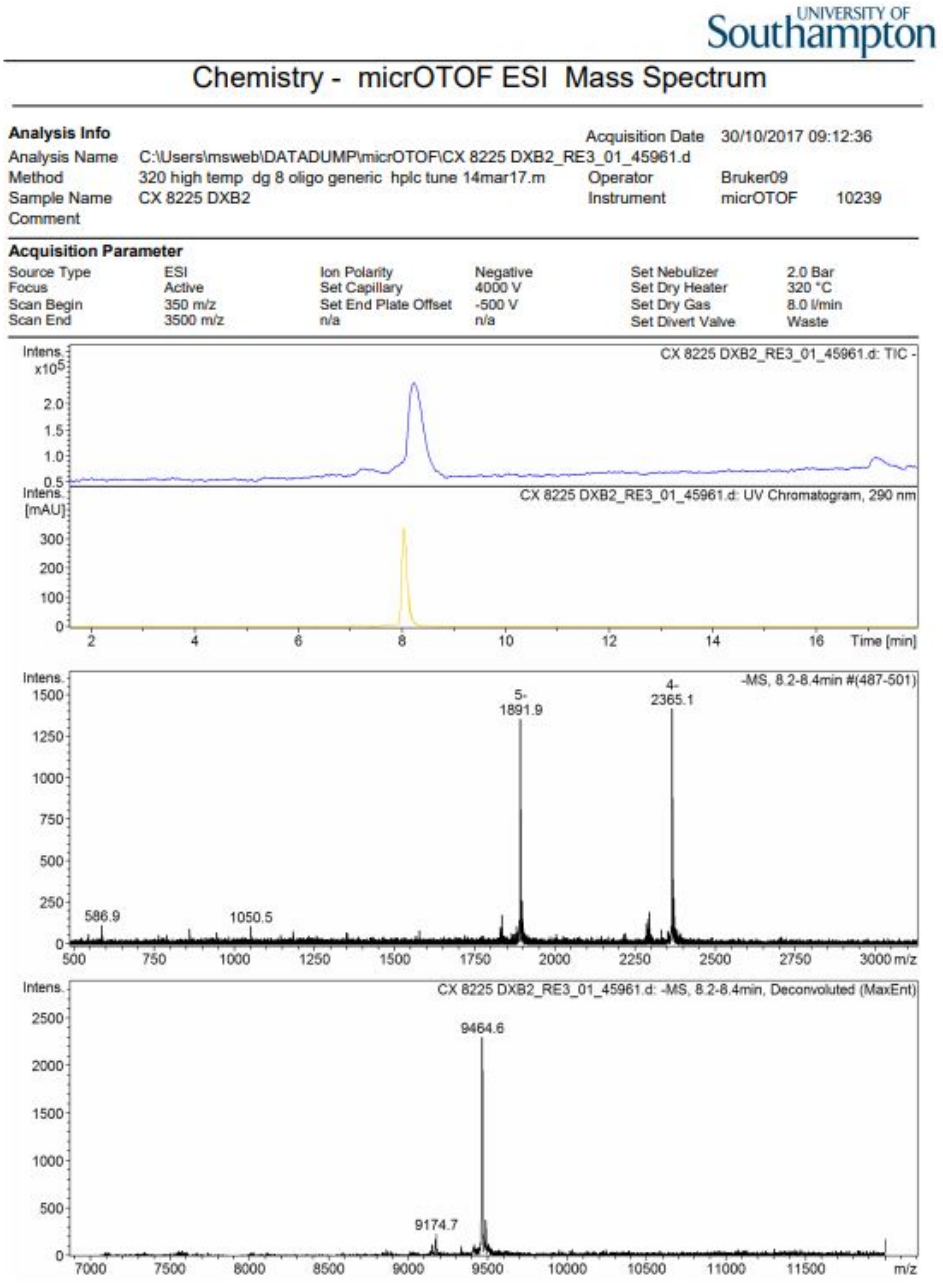

Figure 27: Mass Spectrometry analysis for DX-B2 ssDNA.

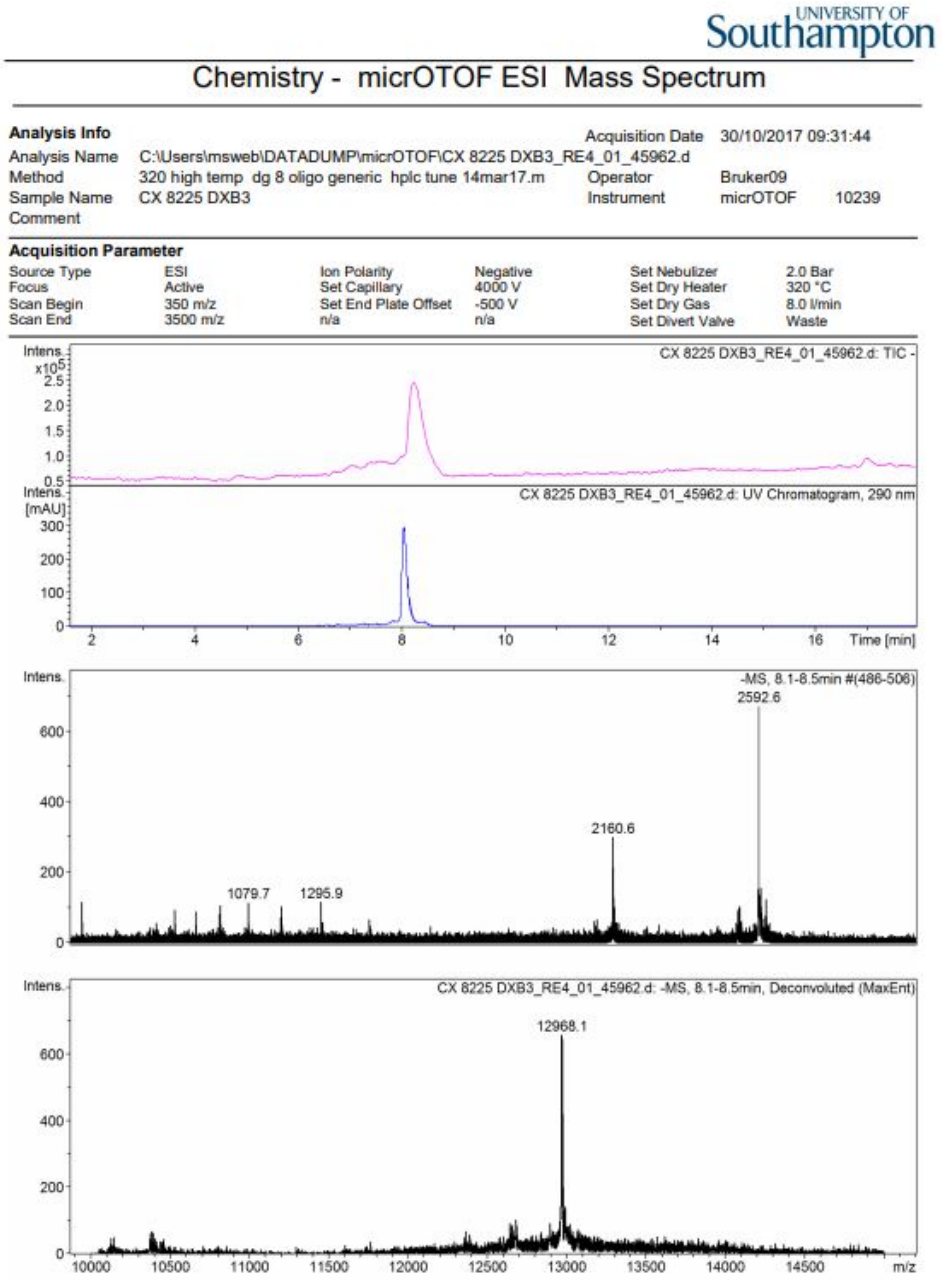

Figure 28: Mass Spectrometry analysis for DX-B3 ssDNA.

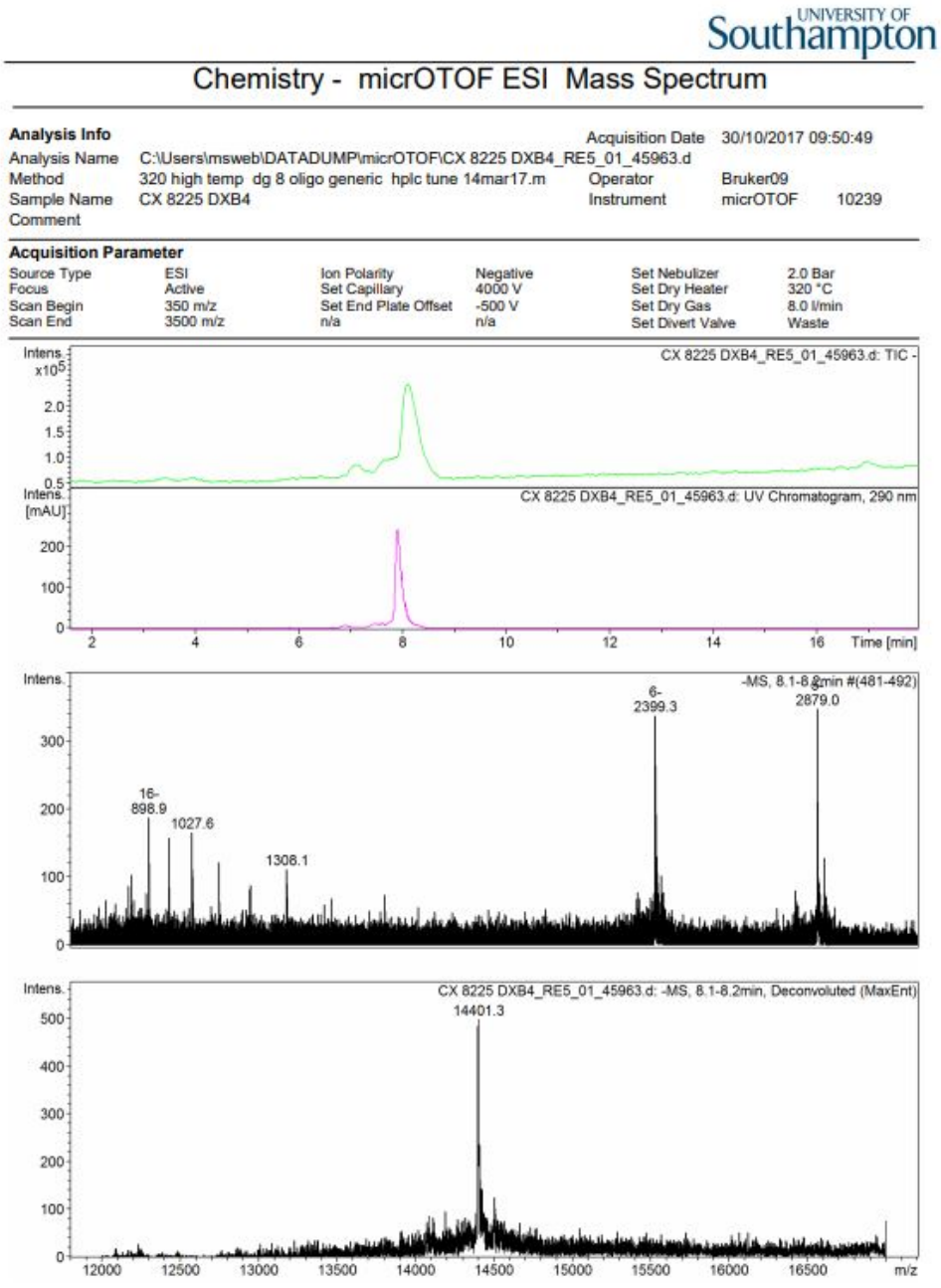

Figure 29: Mass Spectrometry analysis for DX-B4 ssDNA.

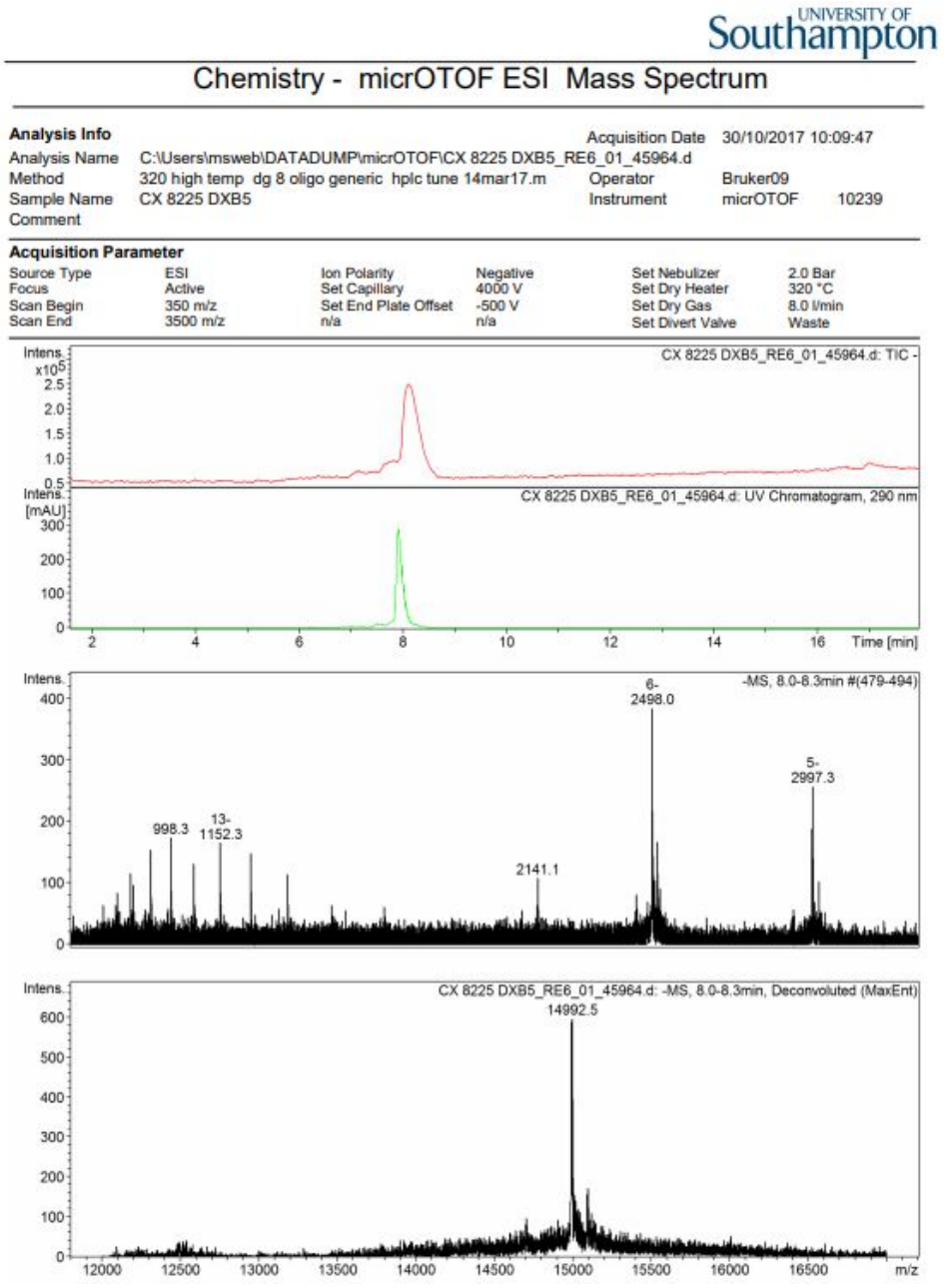

Figure 30: Mass Spectrometry analysis for DX-B5 ssDNA.

Appendix C

Cytochrome b562 Variants

SA - "short axis" variant cytochrome b562 with a p-azido-L-phenylalanine (AzF) mutation at position 5 and a cysteine mutation opposite that at position 104.

E418nm = 52,000

Sequence: ADLE-AzF-NMETLNDNLKVIEKADNAAQVKDALTKMRAAALD-AQKATPPKLEDKSPDSPEMKDFRHGFDILVGQIDDALKLANEG- KVKEAQAAAE-QLKTTRNAYHQCYR

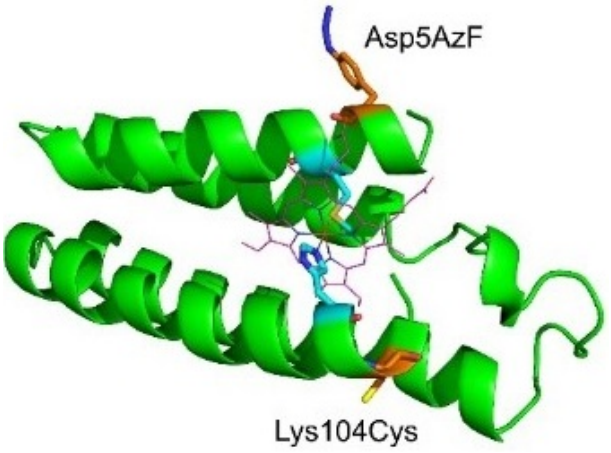

Figure 31: Short axis (SA) cytochrome b_{562} with azF mutation at position 5 and cysteine mutation at position 104.

LA - ”long axis” variant cytochrome b562 with a p-azido-L-phenylalanine (AzF) mutation at position 50 and a cysteine opposite at 21.

E418nm = 85,000.

Sequence: ADLEDNMETLNDNLKVIEKACNAAQVKDALTKMRAAALDAQKA-TPPKLE-AzF-KSPDSPEMKDFRHGFDILVGQIDDALKLANEG- KVKEAQAAAE-QLKTTRNAYHQKYR

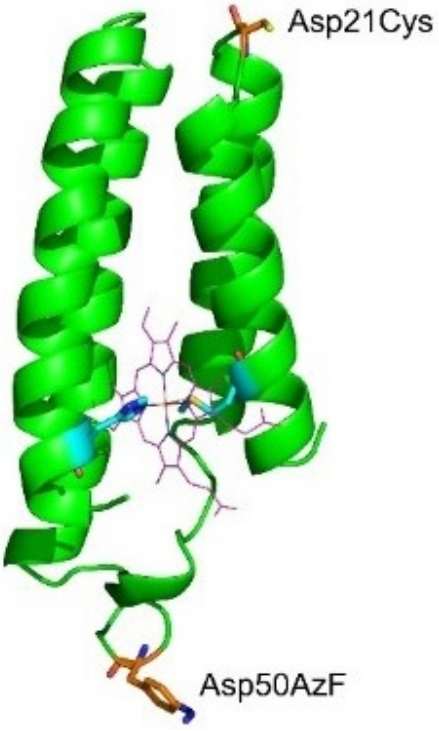

Figure 32: Long axis (LA) cytochrome b_{562} with azF position at 50 and cysteine mutatuion at position 21.