# Manipulating DNA with Magneto-Optical Tweezers and Multiscale Simulation

Jack William Shepherd

Doctor of Philosophy

University of York

Physics

September 2018

# Abstract

Manipulating DNA with optical and magnetic tweezers has been of interest to biophysicists for some time, as the mechanical response of DNA to force and torque is central to its biological function. In this work, methods developed for bespoke transverse magneto-optical tweezers are presented. The tweezers are constructed of a near-infrared laser and surrounding Helmholtz coils for translational and rotational force transduction respectively, and are built such that the DNA under investigation is held transverse to the microscope field of view, which makes this geometry suitable for combined fluorescence and force transduction assays. The persistence length of DNA as measured with this microscope was in accordance with previous work, and undertwisting assays demonstrated that in agreement with literature a limiting force exists between 0.55 and 0.7 pN above which the DNA will not shorten for low values of applied $\sigma$. The intercalating dye YOYO-1 was used to obtain fluorescence images. Hysteresis-like behaviour in over and undertwisted DNA under tension was observed. The severity of buckling for bare DNA and DNA/YOYO-1 complexes were shown to be different by approximately 1.25×.

Molecular simulations of torsionally constrained DNA under overstretching were also performed. oxDNA simulations demonstrated that the stretching perturbation is localised to the moving end of the DNA, and subsequently short fragments of DNA were investigated to understand the perturbed region in atomic detail.

GC rich DNA was shown to be consistently more stable under overstretching with torsional constraint than AT rich DNA. However the response of each construct was determined by specific base pair sequence. A repeating unit of AT was found to form stable periodic structures under overstretching by creating non-canonical hydrogen bonds.

# Contents

**Abstract** **2**

**Table of Contents** **6**

**List of Tables** **7**

**List of Figures** **13**

**Acknowledgements** **14**

**Author's Declaration** **15**

**List of publications arising from this project** **16**

**1 Introduction** **17**

1.1 Structure of this thesis . . . . . . . . . . . . . . . . . . . . . . . . . . . . 17
1.2 DNA . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 18
1.3 Structural properties of DNA . . . . . . . . . . . . . . . . . . . . . . . . 19
1.4 Molecular Machines . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 22
1.4.1 What is a Molecular Machine? . . . . . . . . . . . . . . . . . . . . 22
1.4.2 DNA Regulation, Repair, and Replication . . . . . . . . . . . . 23
1.5 Simulations of DNA . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 24
1.5.1 Atomistic Molecular Dynamics . . . . . . . . . . . . . . . . . . . 24
1.5.2 Coarse-Grained Molecular Dynamics . . . . . . . . . . . . . . . 28
1.5.3 Polymer Physics and Statistical Mechanics . . . . . . . . . . . . 30
1.6 Experimental Approaches to Measuring and Applying Forces . . . . . . 31
1.6.1 Magnetic and Optical Tweezers . . . . . . . . . . . . . . . . . . . 31
1.6.1.1 Pulling on and Twisting Tethers . . . . . . . . . . . . 32
1.6.1.2 Flow- and Electric Field-Based Experiments . . . . . 35
1.6.2 Comparison of Methods . . . . . . . . . . . . . . . . . . . . . . . 36

**2 Experimental Techniques 38**
- 2.1 Target Experimental Geometry . . . 39
- 2.2 DNA Design . . . 40
  - 2.2.1 The 4 kbp Construct . . . 40
  - 2.2.2 The 15 kbp Construct . . . 44
- 2.3 Imaging . . . 48
  - 2.3.1 Optical design . . . 48
    - 2.3.1.1 Design of the fluorescence excitation path . . . 48
    - 2.3.1.2 After excitation: the emission path . . . 51
  - 2.3.2 Choice of fluorescent dyes . . . 53
    - 2.3.2.1 Keeping the lights on . . . 55
    - 2.3.2.2 The Fluorescence Imaging Buffer . . . 56
  - 2.3.3 Data acquisition software and settings . . . 59
- 2.4 Tweezers . . . 60
  - 2.4.1 Principles of optical tweezing . . . 60
    - 2.4.1.1 The large-bead limit . . . 60
    - 2.4.1.2 Practical considerations . . . 63
  - 2.4.2 Design and construction of the optical tweezers . . . 64
  - 2.4.3 Measuring Forces with the QPD . . . 65
  - 2.4.4 Quantifying Trap Stiffness . . . 72
  - 2.4.5 Magnetic Tweezing . . . 73
    - 2.4.5.1 Applying torque with Helmholtz coils . . . 74
- 2.5 Final design . . . 75
  - 2.5.1 Coupling the fluorescence and trapping paths . . . 75
  - 2.5.2 Aligning the optical trap . . . 76
  - 2.5.3 Helmholtz coils and the custom slide holder . . . 77
- 2.6 Selecting magnetic beads . . . 79
- 2.7 Demonstrating trapping and rotating . . . 80
  - 2.7.1 Correcting the focus . . . 83
  - 2.7.2 Force clamping . . . 84
- 2.8 Bead Assays . . . 86
  - 2.8.1 Surface Immobilisation of the antiDIG Beads . . . 86
  - 2.8.2 Conjugation of the DNA Construct . . . 88
  - 2.8.3 Tether Formation . . . 91
- 2.9 Removing Excess Dye . . . 92
- 2.10 Fitting the Wormlike Chain . . . 95

2.11 Experiment Protocol . . . 96
2.12 Summary . . . 98

**3 DNA Manipulation *in vitro* 100**
3.1 Brightfield Assays . . . 100
    3.1.1 Force-Extension Behaviour . . . 100
    3.1.2 Force Clamp Assays . . . 107
3.2 DNA/YOYO-1 Complexes . . . 112
    3.2.1 Force-Extension Behaviour . . . 112
    3.2.2 Twist-Extension Relations . . . 114
    3.2.3 Fluorescence Imaging with YOYO-1 . . . 118
3.3 Summary . . . 120

**4 A Brief Guide to Molecular Simulation 123**
4.1 Why Perform Multiscale Modelling on DNA? . . . 123
4.2 The Models . . . 124
    4.2.1 Amber . . . 124
    4.2.2 oxDNA . . . 127
4.3 The Langevin Thermostat . . . 129
4.4 Implicit Solvation . . . 130
    4.4.1 The Generalised Born Solvation Model . . . 130
        4.4.1.1 Ionic Strength . . . 131
    4.4.2 Solvation in oxDNA . . . 132
4.5 Applying External Force and Torque . . . 132
4.6 Minimising the Potential Energy of the System . . . 135
4.7 DNA Structures and Production MD . . . 136
4.8 Analysis of Trajectories . . . 137
4.9 Comparison to Experiment . . . 138
4.10 Summary . . . 138

**5 Simulating DNA 139**
5.1 Simulation Design . . . 139
5.2 oxDNA Simulations . . . 141
    5.2.1 poly-d(A)·poly-d(T) . . . 141
    5.2.2 poly-d(C)·poly-d(G) . . . 145
5.3 Amber Simulations . . . 147
    5.3.1 Minimisation . . . 147

5.3.1.1 Equilibration . . . . . . . . . . . . . . . . . . . . . . . 151
5.3.2 poly-d(A)·poly-d(T) . . . . . . . . . . . . . . . . . . . . . . . . 153
5.3.3 poly-d(AT)·poly-d(AT) . . . . . . . . . . . . . . . . . . . . . . 160
5.3.4 poly-d(C)·poly-d(G) . . . . . . . . . . . . . . . . . . . . . . . . 168
5.3.5 poly-d(CG)·poly-d(CG) . . . . . . . . . . . . . . . . . . . . . . 173
5.4 Summary . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 177

**6 Discussion** **180**
6.1 Future Work . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 183

**A Protocol for Constructing the 4 kbp Tether** **186**

**B Heatmaps for 100 ps/Å Replicas** **191**
B.1 poly-d(A)·poly-d(T) . . . . . . . . . . . . . . . . . . . . . . . . . 192
B.2 poly-d(AT)·poly-d(AT) . . . . . . . . . . . . . . . . . . . . . . . 194
B.3 poly-d(C)·poly-d(G) . . . . . . . . . . . . . . . . . . . . . . . . . 195
B.4 poly-d(CG)·poly-d(CG) . . . . . . . . . . . . . . . . . . . . . . . 197

**Bibliography** **214**

# List of Tables

1.1 Key structural properties of A-DNA, B-DNA, and Z-DNA, the three most common forms of double helix DNA . . . . . . . . . . . . . . . . . 20

2.1 Physical properties of the chosen magnetic bead . . . . . . . . . . . . . 80

5.1 Heatmap coding for the canonical bond structure heatmaps . . . . . . 156
5.2 Heatmap coding for the non-canonical heatmaps . . . . . . . . . . . . 157
5.3 Heatmap coding for the full colour hydrogen bond heatmaps. Columns are canonical hydrogen bonds, rows are non-canonical hydrogen bonds 157

# List of Figures

1.1 The structure of canonical A- and B-form DNA . . . . . . . . . . . . . 18

1.2 Diagram of nucleotide geometry with base and deoxyribose sugar ring indicated . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 19

1.3 a) Stereoscopic projection of the structure of S-DNA found through molecular dynamics. The helix has unwound leaving the characteristic “ladder-like” structure, with inclined base-pairs joined through canonical hydrogen bonds. Drawing reproduced from Lohikoski *et al. [1]* b) Structure of P-DNA found from experiment and constructed using a space-filling algorithm. The backbone, represented in pink, has become the central part of the double helix, with the bases (blue and yellow) turned outwards and exposed. This drawing was taken from Allemand *et al.* [2] . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 21

1.4 A Cartesian representation of a 2D freely-jointed chain polymer consisting of 10 Kuhn sections . . . . . . . . . . . . . . . . . . . . . . . . . 30

2.1 The target experiment. A 5 $\mu$m antiDIG functionalised bead is immobilised on the cover slip surface. A 3 $\mu$m streptavidin-coated bead is trapped optically. The target molecule of interest forms a horizontal tether between the two beads. By applying force with the optical trap or torque using bespoke Helmholtz coils mounted on the microscope (not pictured), the tethered molecule can be directly manipulated. As the tether is transverse to the objective lens, the full contour of the molecule is visible. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 39

2.2 Schematic diagram of the PCR and restriction undertaken to create the 500 bp handles for the 4 kbp construct . . . . . . . . . . . . . . . . . . 41

2.3 PCR schematic for the midsection of the 4 kbp construct . . . . . . . . 42

2.4 The 4 kbp construct tested with agarose gel electrophoresis . . . . . . . 43

2.5 Agarose gels demonstrating the construction and activity of the 15 kbp tether . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 47
2.6 Excitation profile of YOYO-1 and the spectral properties of the corresponding excitation beam . . . . . . . . . . . . . . . . . . . . . . . . . . . 49
2.7 Schematic drawing of the blue fluorescence path of the microscope . . . 50
2.8 The emission spectrum of YOYO-1 and the transmission properties of the relevant region of the emission filter . . . . . . . . . . . . . . . . . . 51
2.9 The microscope emission path. During a trapping experiment, the light initially passes through the focus correcting module. Following that the light passes through a 75 mm lens. A flip mirror (marked inside a dashed box) selects the Andor iXon+ cameras or the Watec webcam which has a larger field of view. If the Andor cameras are selected, the light passes into a 150 mm lens, expanding the beam by 2X, and subsequently an Andor Tucam module which splits the light according to wavelength and focusses the red light on to one camera and the green light on to another, and in this way two colour imaging is possible. If the Watec webcam is selected, the light meets another 75 mm lens and the beam is therefore not expanded. The optics within the microscope body itself including the dichroic mirror and tube lens are not shown. . . . . . . . . . . . . 52
2.10 Three possible dye binding modes . . . . . . . . . . . . . . . . . . . . 54
2.11 Average intensity per 40 ms exposure through time for a) DNA+YOYO-1 without the oxygen scavenging system and b) with the oxygen scavenging system present . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 57
2.12 Average intensity per 40 ms exposure through time for DNA+YOYO-1 in an oxygen scavenger without Trolox . . . . . . . . . . . . . . . . . . 58
2.13 Three possible beads positions in an optical trap and the resultant forces 62
2.14 Absorption spectrum of water . . . . . . . . . . . . . . . . . . . . . . 63
2.15 Schematic of the optical path for the laser tweezers . . . . . . . . . . . 65
2.16 Stages of the data analysis to generate trap stiffness . . . . . . . . . . . 67
2.17 Theoretical QPD response to an object scanning through the centre . . 70
2.18 QPD response as a bead is scanned along the $y$ axis in the centre of the trap. The central region has been fitted to a straight line using the `curve_fit` routine in SciPy. . . . . . . . . . . . . . . . . . . . . . . . 71
2.19 Trap stiffness as a function of height . . . . . . . . . . . . . . . . . . 73
2.20 Images of nIR light reflected from a water sample . . . . . . . . . . . 77
2.21 Photograph of the custom Helmholtz coils . . . . . . . . . . . . . . . 78

2.22 A trapped bead 1 $\mu$m above the cover slip . . . . . . . . . . . . . . . . 81
2.23 Fluorescence images showing a trapped bead rotating . . . . . . . . . . 81
2.24 Kymograph of the optically trapped, rotating bead . . . . . . . . . . . 82
2.25 The shift in imaging plane due to the corrective module in the imaging path . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 82
2.26 The effect of the corrective module on the appearance of a bead optically trapped 1 $\mu$m above the cover slip . . . . . . . . . . . . . . . . . . . . 84
2.27 The response through time of the force clamp to sudden application of a 2 pN target force . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 85
2.28 Mean number of immobilised 5 $\mu$m antiDIG beads per field of view on the benchtop microscope in different ionic conditions . . . . . . . . . . 88
2.29 Subsequent frames from a fluorescence acquisition of the 15 kbp construct conjugated to an antiDIG bead . . . . . . . . . . . . . . . . . . . . . . 89
2.30 Total fluorescence intensity for antiDIG beads through time . . . . . . 90
2.31 Possible tethering geometries for the 1 $\mu$m DNA tether, drawn to scale 91
2.32 Bar charts showing a) integrated fluorescence intensity of beads in the presence of YOYO-1 and YOYO-1/DNA complexes, and b) intensity of beads with YOYO-1 and YOYO-1/DNA complexes after size exclusion column purification . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 93
2.33 Fluorescent images of DNA/YOYO-1 purified with a size exclusion column 95
2.34 Stiffness of trap for 83 magnetic beads approximately 2.5 $\mu$m from the cover slip surface, fitted to a Gaussian distribution. . . . . . . . . . . . 97

3.1 Triangular movement of the nanostage for generating force-extension curves and the resulting force applied . . . . . . . . . . . . . . . . . . . 101
3.2 Four wormlike chain fits to uncorrected force-extension curves . . . . . 102
3.3 Example $q$ angle as a function of displacement . . . . . . . . . . . . . . 103
3.4 A force-extension curve corrected with force inference . . . . . . . . . . 104
3.5 A force-extension curve corrected with the absolute force method with the mean force at low extension subtracted . . . . . . . . . . . . . . . . 105
3.6 Twist-extension curves for a single DNA tether under low (0.55 pN) force 108
3.7 Force through time for a DNA tether undertwisted and overtwisted with the force clamp set to 0.55 pN . . . . . . . . . . . . . . . . . . . . . . . 109
3.8 The twist-extension behaviour for a single dsDNA molecule under 1.1 pN of tension . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 110
3.9 Force extension curves for DNA complexed with YOYO-1 . . . . . . . . 113

3.10 Extension with overtwisting a DNA/YOYO-1 complex with applied 2 pN tension . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 114
3.11 Force vs applied relative twist for the DNA/YOYO-1 complex in Figure 3.10 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 115
3.12 Extension vs applied relative twist for a DNA/YOYO-1 complex with applied tension of 0.7 pN . . . . . . . . . . . . . . . . . . . . . . . . . 116
3.13 Comparison of gradients before and after the buckling points for DNA with and without YOYO-1 . . . . . . . . . . . . . . . . . . . . . . . . . 117
3.14 Fluorescence image of a tether snapping under illumination . . . . . . . 118
3.15 DNA tether imaged with 1 mW power in oxygen scavenging buffer . . . 119

4.1 Magnifying a DNA experiment with successive levels of molecular simulation . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 124
4.2 The oxDNA model with interactions for the left middle base indicated . 127
4.3 The stretching procedure for simulations in this work . . . . . . . . . . 134

5.1 Total, potential, and kinetic energy over 3 ns for poly-d(A)·poly-d(T) stretched by 21.3 nm simulated in oxDNA . . . . . . . . . . . . . . . . 141
5.2 Total, potential, and kinetic energies for poly-d(A)·poly-d(T) stretched using oxDNA . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 142
5.3 Simulated conformations of poly-d(A)·poly-d(T) simulated in oxDNA . 143
5.4 Total, kinetic, and potential energy for poly-d(C)·poly-d(G) over 3 ns when stretched by 20 nm in oxDNA . . . . . . . . . . . . . . . . . . . 144
5.5 Total, potential, and kinetic energies over a full stretching simulation of poly-d(C)·poly-d(G) in oxDNA . . . . . . . . . . . . . . . . . . . . . 145
5.6 poly-d(C)·poly-d(G) stretched by a) 0 nm; b) 10 nm; c)20 nm in an oxDNA simulation . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 146
5.7 0K minimisation of 24 bp dsDNA in Amber after application of a twist to one terminal base pair . . . . . . . . . . . . . . . . . . . . . . . . . 147
5.8 Energy during a 0K minimisation for a) poly-d(A)·poly-d(T) and b) poly-d(C)·poly-d(G) stretched by 1 Å in Amber . . . . . . . . . . . . . . . 149
5.9 Minimisation with $10^6$ steepest descent steps followed by conjugate-gradient for a) poly-d(A)·poly-d(T) and b) poly-d(C)·poly-d(G) stretched in Amber by 30 Å. . . . . . . . . . . . . . . . . . . . . . . . . . . . . 150
5.10 Temperature and energy over a 500 ps simulation of poly-d(A)·poly-d(T) stretched by 1 Å . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 152

5.11 Temperature and energy over a 500 ps simulation of poly-d(A)·poly-d(T) stretched by 30 Å . . . . . . . . . . . . . . . . . . . . . . . . . . . . 152
5.12 Structures from the stretching simulation of untwisted poly-d(A)·poly-d(T) 153
5.13 Hydrogen bonds for stretched and twisted poly-d(A)·poly-d(T) in Amber 155
5.14 Heatmaps showing the hydrogen bonding of poly-d(A)·poly-d(T) with an applied stretch of 25 Å and $\sigma = -0.068$ . . . . . . . . . . . . . . . . 157
5.15 Structure heatmaps for hydrogen bonds for poly-d(A)·poly-d(T) stretched 25 Å and with $\sigma = 0$ . . . . . . . . . . . . . . . . . . . . . . . . . . . . 158
5.16 Structural heatmaps showing canonical hydrogen bonds, non-canonical hydrogen bonds, and all hydrogen bonds for poly-d(A)·poly-d(T) stretched by 25 Å with $\sigma = 0.068$ . . . . . . . . . . . . . . . . . . . . . . . . . . 159
5.17 van der Waals energy through time for poly-d(A)·poly-d(T) over a simulation stretching it to 30 Å with $\sigma = 0, \pm0.068$ with a 100 point rolling average applied . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 159
5.18 Upper: starting structure for the 15 ns poly-d(AT)·poly-d(AT) simulation with $\sigma = 0$. Lower: structure at a stretch of 30 Å . . . . . . . . . . 160
5.19 Hydrogen bond heatmaps of canonical and non-canonical hydrogen bonds for poly-d(AT)·poly-d(AT), with $\sigma = 0, \pm0.068$. Once again the $y$ axis represents the simulated time while the horizontal axis shows the number of the base pair. An instantaneous stretch of 1 Å is applied every 500 ps for the $\sigma = \pm0.068$ simulations and every 1 ns for the $\sigma = 0$ simulation. 162
5.20 Hydrogen bond heatmaps for 100 ps/Å replica of poly-d(AT)·poly-d(AT) 163
5.21 Canonical, non-canonical, and all hydrogen bonds for poly-d(AT)·poly-d(AT) with $\sigma = -0.068$ stretched 25 Å . . . . . . . . . . . . . . . . . . 164
5.22 Canonical, non-canonical and all hydrogen bonds for untwisted poly-d(AT)·poly-d(AT) stretched by 25 Å . . . . . . . . . . . . . . . . . . . 164
5.23 Canonical, non-canonical and all hydrogen bonds for poly-d(AT)·poly-d(AT) with $\sigma = 0.068$ stretched by 25 Å . . . . . . . . . . . . . . . . . . 165
5.24 VdW energies for poly-d(AT)·poly-d(AT) with $\sigma = 0, \pm0.068$. A 100 point rolling average has been applied. . . . . . . . . . . . . . . . . . . 165
5.25 Stacking interactions and hydrogen bonding for poly-d(AT)·poly-d(AT) with $\sigma = -0.068$ . . . . . . . . . . . . . . . . . . . . . . . . . . . . 167
5.26 poly-d(C)·poly-d(G) with $\sigma = 0$ stretched by 0 (upper) and (lower) 30 Å 168

5.27 Heatmaps of canonical and non-canonical hydrogen bonds for poly-d(C)·poly-d(G) with $\sigma = 0, \pm 0.068$. On the horizontal axes is the base pair, and on the vertical axis is the simulated time. A stretch of 1 Å is applied in a single step every 500 ps for the $\sigma = \pm 0.068$ simulations and every 1 ns for the $\sigma = 0$ simulation. . . . . . . . . . . . . . . . . . . . . . . . . . 169
5.28 Canonical, non-canonical, and all hydrogen bonds for poly-d(C)·poly-d(G) with $\sigma = -0.068$ and extension 25 Å . . . . . . . . . . . . . . . . . 170
5.29 Canonical, non-canonical and all hydrogen bonds for untwisted poly-d(C)·poly-d(G) at a stretch of 25 Å . . . . . . . . . . . . . . . . . . . . . 171
5.30 Canonical, non-canonical and all hydrogen bonds for poly-d(C)·poly-d(G) with $\sigma = 0.068$ stretched by 25 Å . . . . . . . . . . . . . . . . . . 171
5.31 Total VdW energy for poly-d(C)·poly-d(G) with $\sigma = 0, \pm 0.068$ . . . . . 172
5.32 Initial (upper) and 30 Å stretched (lower) structures of poly-d(CG)·poly-d(CG) with $\sigma = 0$ . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 173
5.33 Heatmaps of canonical and non-canonical hydrogen bonds for stretched and twisted poly-d(CG)·poly-d(CG) with $\sigma = 0, \pm 0.068$. The horizontal axis shows the base pair, and along the vertical axis is the simulated time. A sharp instantaneous stretch of 1 Å is applied every 500 ps for the $\sigma = \pm 0.068$ simulations and every 1 ns for the $\sigma = 0$ simulation. . 174
5.34 Canonical, non-canonical and all hydrogen bonds for poly-d(CG)·poly-d(CG) with $\sigma = -0.068$ and applied stretch 25 Å . . . . . . . . . . . . 175
5.35 Watson-Crick, non-Watson-Crick, and all hydrogen bonds for poly-d(CG)·poly-d(CG) with $\sigma = 0$ and applied stretch 25 Å . . . . . . . . . . . . . . . . 176
5.36 Canonical, non-canonical and all hydrogen bonds for poly-d(CG)·poly-d(CG) with $\sigma = 0.068$ and applied stretch 25 Å . . . . . . . . . . . . . 176
5.37 Total van der Waals energy through time for poly-d(CG)·poly-d(CG) with $\sigma = 0, \pm 0.068$ . . . . . . . . . . . . . . . . . . . . . . . . . . . 177

B.1 poly-d(A)·poly-d(T) 100 ps/Å replica 1 . . . . . . . . . . . . . . . . . . . 192
B.2 poly-d(A)·poly-d(T) 100 ps/Å replica 2 . . . . . . . . . . . . . . . . . . . 193
B.3 poly-d(AT)·poly-d(AT) 100 ps/Å replica 2 . . . . . . . . . . . . . . . . 194
B.4 poly-d(C)·poly-d(G) 100 ps/Å replica 1 . . . . . . . . . . . . . . . . . . . 195
B.5 poly-d(C)·poly-d(G) 100 ps/Å replica 2 . . . . . . . . . . . . . . . . . . . 196
B.6 poly-d(CG)·poly-d(CG) 100 ps/Å replica 1 . . . . . . . . . . . . . . . . . 197
B.7 poly-d(CG)·poly-d(CG) 100 ps/Å replica 2 . . . . . . . . . . . . . . . . . 198

# Acknowledgements

I would first like to thank my supervisors: Mark Leake, Matt Probert and Robert Greenall. Their thoughtful suggestions, patience, and help have been invaluable and I am grateful that I had each supervisor on board.

Adam Wollman, Laurence Wilson, Steve Quinn, and Phil Hasnip are all great company and profoundly knowledgeable. I can't thank them enough for the many constructive talks and corridor advice they have given me.

This has been an extremely collaborative project and I would like also to especially thank Zhaokun (Jack) Zhou, who I worked primarily with. His depth of knowledge and boundless enthusiasm were a joy to behold; a better collaborator cannot be imagined. I would also like to thank Helen Miller for a lot of instruction at the outset of my PhD and for the many invaluable tips along the way, and Jamie Howard for his help making the DNA that made all this work possible.

I would like also to thank all my friends in York. There are too many to list but I would like to highlight Aaron Hopkinson for the many robust discussions we enjoyed late at night, Ed Higgins for being an incredibly good sport, and Jacob Wilkins for a depth of knowledge about bash which is surely unmatched anywhere in the world. Those three were the first people I really got to know in York and without them it would not have been anything like as good an experience. Others who have made this time worthwhile include Sviatlana Shashkova, James Ramsden, Erik Hedlund, and Neville Yee.

Finally, thanks to my family. Without their moral (and often financial) support I would never have gone this far.

# Author's Declaration

I declare that this thesis is a presentation of original work and I am the sole author except for the specific collaborations detailed below. This work has not previously been submitted for a degree at this or any other university. All sources are acknowledged as references.

- The 4kbp construct was designed and made by Zhaokun Zhou
- The 15 kbp tether was made jointly with Jamieson Howard
- The trapping optics were originally designed by Zhoukun Zhou
- The Helmholtz coils and holders were designed, fabricated, and fitted by Zhaokun Zhou
- The excitation and emission optics were designed and built by Helen Miller
- The trap stiffness with height measurement in Chapter 2 was performed and analysed jointly with Zhaokun Zhou
- Zhaokun Zhou wrote the code to acquire data and run the force clamp

# List of publications arising from this project

## Peer-reviewed publications

1. Miller, H., Zhou, Z., Shepherd, J., Wollman, A.J. and Leake, M.C., 2017. Single-molecule techniques in biophysics: a review of the progress in methods and applications. Reports on Progress in Physics, 81(2), p.024601.

## Unreviewed Conference Abstracts

1. Shepherd, J., Zhou, Z., Greenall, R., Probert, M. and Leake, M., 2017. Transverse fluorescence microscopy with magnetic and optical tweezers. Biophysical Journal, 112(3), p.299a.

## Poster Presentations

1. Shepherd, J, Zhou, Z, Greenall, R., Probert, M. and Leake, M., 2016. Transverse Optical and Magnetic Tweezers for DNA, Presented at the British Biophysical Society
2. Shepherd, J, Zhou, Z, Greenall, R., Probert, M. and Leake, M., 2017. Transverse fluorescence microscopy with magnetic and optical tweezers. Presented at the Biophysical Society
3. Shepherd, J, Zhou, Z, Greenall, R., Probert, M. and Leake, M., 2017. Transverse fluorescence microscopy with magnetic and optical tweezers. Presented at the International Union of Pure and Applied Biophysics

# Chapter 1

# Introduction

## 1.1 Structure of this thesis

This thesis describes methods developed for a novel combined magneto-optical tweezing tool consisting of an optical trap and Helmholtz coils mounted to the microscope stage, and its subsequent usage. Complementary multiscale molecular dynamics simulations which explore the conformation of DNA under over-stretched, rotationally constrained conditions were also performed, and the design of these simulations and their results are presented.

In this Chapter, the motivation for building such an implement is given, and the previous work in the field is summarised. The various methods of applying force and torque to DNA are discussed and the advantages and disadvantages of each are explained. Previous molecular dynamics work on DNA under tension and torsion are also explored. In Chapter 2, the physical principles of optical and magnetic tweezers are described along with practical considerations during their construction, their final design, and methods used to characterise trap strength and extract force information. The method development undertaken is also presented, and in Chapter 3 the results of the subsequent experiments are presented. Chapter 4 gives an overview of the physics, mathematics, and design considerations of both atomistic and coarse-grained molecular dynamics simulation. The molecular dynamics simulation design, results, and hydrogen bond analysis constitute Chapter 5, and in the final Chapter the results of this work are discussed and future perspectives considered.

## 1.2 DNA

The double helical structure of double-stranded deoxyribonucleic acid dsDNA [3], shown in Figure 1.1, and its information carrying function are two of the most ubiquitous scientific ideas in popular culture. What is often less realised is that DNA remains a target for ongoing research in institutions throughout the world. The central dogma of molecular biology states that it carries a code which may be transcribed to produce proteins and hence control the cell, but this is only the start of its many functional properties. The mechanical properties of DNA – how it stretches, coils, interacts with ligands – are as vital to its biological importance as the instruction set it carries: indeed, many proteins are encoded more than once in the genome and which (if any) is expressed is a function of the wider context of the DNA and cell.

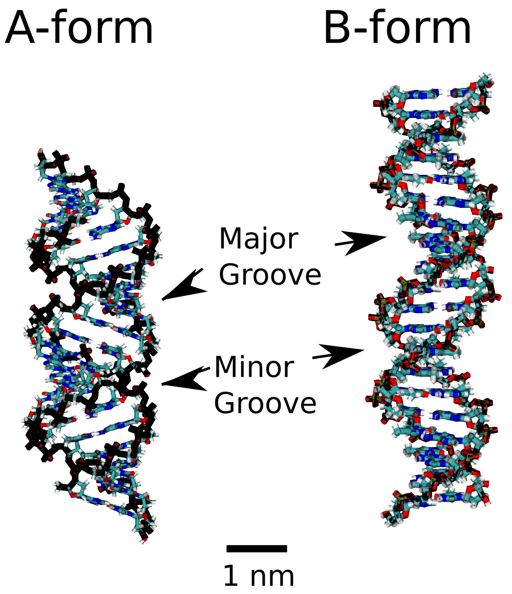


Figure 1.1: The structure of canonical A- and B-form DNA

At its most basic level, DNA is a polymer composed of a combination of four monomers: adenine, thymine, cytosine, and guanine, collectively referred to as the nucleotides. Each is composed of a nucleic base attached to a sugar ring as seen in Figure 1.2. The deoxyribose sugar rings may be linked by a directional phosphate backbone to form a strand of nucleotides – i.e. single-stranded DNA (ssDNA) – and two compatible strands may then anneal to form a double strand of DNA. Through stacking interactions between successive base pairs on each strand, this double stranded DNA will preferentially fold to form a stable right-handed double helix. If the phosphate backbone is sufficiently electrostatically screened from itself and the DNA is put under torsion, it will wind around itself again to form tertiary structures such as plectonemes [4, 5], and can further condense into chromosomes [6, 7]. In this way DNA may be so tightly packaged that the entire human genome – approximately two metres in length – fits inside the nucleus of the cell – typically 2-10 $\mu$m.

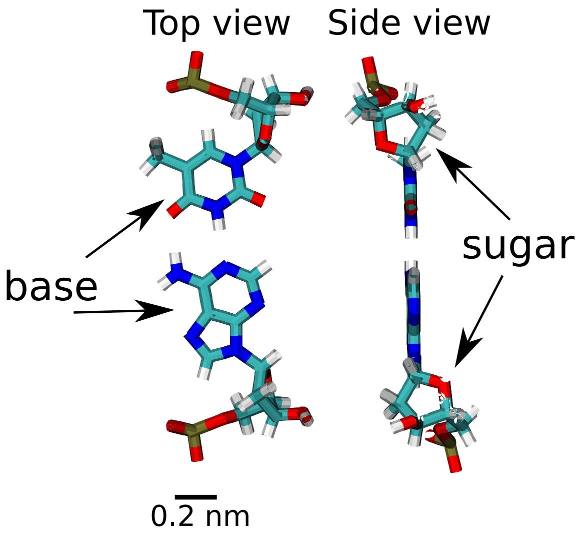


Figure 1.2: Diagram of the nucleotide geometry with base and deoxyribose sugar ring indicated

## 1.3 Structural properties of DNA

To specify a DNA conformation, multiple different parameters must be taken into account. The inter-base-pair properties are those which arise between successive base-pairs. The two key inter-base-pair parameters for determining conformation are the distance along the helical axis between successive base pairs which is known as the *rise* of the structure, and the angular offset between the two successive base pairs which is called the DNA's *twist*. Intra-base-pair parameters are those which occur within the base-pair itself, and include the relative displacement of the bases and their relative orientation, and this angular difference in orientation gives rise to larger-scale conformational parameters which are of key importance – the major and minor grooves. The asymmetry in the base orientations leads to asymmetry in the distance between successive phosphate backbones when looking at the double helix as a whole. Within the smaller minor groove are the sugar rings of the nucleotides, while at the bottom of the major groove are the nucleotides' bases themselves.

Using these parameters, there are multiple distinct conformations that DNA can assume under the right conditions, and these are are generally given letter names. The two most common DNA conformations are known as the A form and B form, and were discovered in the early days of DNA crystallography. Both of these conformations are right handed, meaning that successive base pairs rotate in a clockwise sense as the helical axis is traversed. A-DNA has been found to be a largely crystallographic structure, while the B-form is the usual conformation found *in vivo*. The two differ in multiple ways. The rise per base-pair is larger in B-DNA than in A-DNA, and the twist is lower. Taken together, this means that A-DNA is a more compact form than the B form, and the major groove is deeper but narrower in the A-form. The A- and B-form are shown in Figure with the major and minor grooves indicated.

More exotic structures were subsequently found, including the Z-form. This is distinct in that it is considerably more extended than either the A-form or B-form and, more unusually, it is left-handed. Its structure has biological implications, with specific proteins existing which bind to Z-DNA but not to B-DNA [8, 9], and Z-DNA regions coexist with B-DNA with a distinct interface [10]. Z-DNA formation is affected by sequence and methylation, and by the presence of negative supercoiling [11]. Together, the A form, B form, and Z form are the three most common conformations of DNA, and their key properties are summarised in Table 1.1.

| | A-DNA | B-DNA | Z-DNA |
|---|---|---|---|
| Chirality | Right-handed | Right-handed | Left-handed |
| Rise per bp (Å) | 2.3 | 3.3 | 3.8 |
| Twist (°) | 32.7 | 34.3 | 30 |
| Base-pairs per helical pitch | 11 | 10.5 | 12 |
| Helical pitch (Å) | 28.2 | 33.2 | 45.6 |

Table 1.1: Key structural properties of A-DNA, B-DNA, and Z-DNA, the three most common forms of double helix DNA

More mechanical perturbation leads to even more exotic DNA structure. Of these, the two which are most relevant to the work in this thesis are the S-form [1] and P-form [2, 12], the structures of which can be seen in Figure 1.3.

S-DNA, also known as S-ladder DNA, arises when DNA is put under external tension with ends free to rotate and relieve torsion. The DNA extends and the helix unwinds such that the DNA appears similar to a ladder with the base pairs forming the “rungs”, though these are highly inclined [1]. The hydrogen bonding within the base-pairs remains intact, the DNA being in effect unwrapped without novel bonding behaviour in evidence.

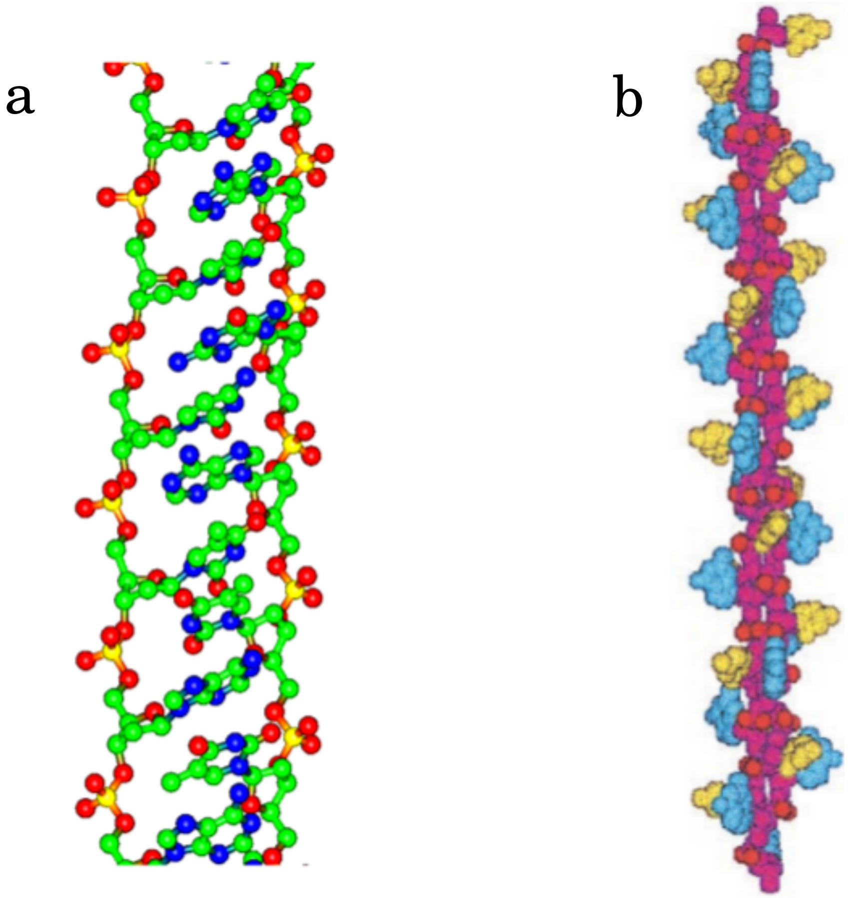


Figure 1.3: a) Stereoscopic projection of the structure of S-DNA found through molecular dynamics. The helix has unwound leaving the characteristic “ladder-like” structure, with inclined base-pairs joined through canonical hydrogen bonds. Drawing reproduced from Lohikoski *et al.* *[1]* b) Structure of P-DNA found from experiment and constructed using a space-filling algorithm. The backbone, represented in pink, has become the central part of the double helix, with the bases (blue and yellow) turned outwards and exposed. This drawing was taken from Allemand *et al.* [2]

P-DNA, so called because it appears similar to a structure for DNA first proposed by Linus Pauling [2], appears in DNA which is under torsional stress as well as tension. Unlike the S-form, base-pairs in this conformation lose their hydrogen bonds and the bases themselves flip outwards such that it is the bases which are exposed and the phosphate backbones are interwound directly [2]. This form of DNA has a very short helical pitch, and appears in distinct segments in overtwisted DNA, which has the effect of preserving more canonical B-DNA regions [12].

Both P-DNA and S-DNA appear in mechanically perturbed systems and therefore may exist *in vivo* at sites which are under manipulation by molecular machines. However, the biological function of these forms – if any – remains unclear.

DNA need not assume a double-helical two-stranded conformation. Triplex DNA has been found, in which a third strand binds to the major groove of an existing double helix [13]. More exotically, a single strand with a high repeating G content can form a quadruplex through G-G interactions [14]. Here, four G bases interact to form a square structure, and if the DNA sequence is such to allow multiples of these to form in succession, an extended structure can be formed, consisting usually of four base quadruplexes. These G-quadruplexes are stable at physiological pH and temperature, and therefore may form *in vivo* and play a biological role.

Clearly, however, this is not the end of the story. Having been formed, wound, and packaged, the DNA still needs to play an important part in the regulation of the cell - activating and deactivating genes through mechanics [15] or methylation [16], expressing proteins, replication, and triggering cell death, and many of these processes are affected by molecular crowding and the proteins bound to the DNA at the time [17]. This is achieved through interactions with other cellular biomolecules, which in the crowded cellular environment is a complex task often requiring and resulting in structural change [18]. For example, deactivation of a gene may be accomplished by binding a ligand to the relevant DNA sequence and thus blocking its transcription [19], or the DNA itself may be put in a conformation that increases or decreases synthesis [20]. In either case, the ligand or DNA has to perform a specific function while many other biomolecules are present and performing their own functions. Similarly, for DNA to replicate it must be first unpacked and unwound into its component single strands even while ligands are binding or unbinding - a process which manifestly requires manipulation of the DNA by an external actor. In general, cellular processes often require direct force transduction, whether that leads to cargo transport [21] or topological changes within the DNA [22], and for this cells make use of molecular machines.

## 1.4 Molecular Machines

### 1.4.1 What is a Molecular Machine?

Like all machines, molecular machines perform mechanical work on targets, but operate at molecular (typically nanometre) length-scales. The conformational changes required for this work to be done are necessarily reversible, but molecular machines have evolved to go through several distinct stages in their operation to prevent repeating the same steps. The molecular machine will start in a relatively high potential energy state, and after getting over an energy barrier will find itself in a lower potential energy state, and therefore will not preferentially return to the previous state due to the high energy barrier, though it is possible. As energy is added to the system, the machine recovers a higher energy state without going backwards, and returns eventually to its original conformation. One cycle of its operation has thus occurred, and it is ready to begin again. Although molecular machines are ubiquitous in the cell and many classes of them exist, here only two key types of molecular machines which directly affect DNA conformation will be discussed.

### 1.4.2 DNA Regulation, Repair, and Replication

Many molecular machines exist within the cell to assist with and perform DNA functions. Key to those, found in all forms of life and consequently well studied, are the topoisomerases. There are two types of topoisomerase, type I (Topo I) and type II (Topo II) [23]. Type I topoisomerases separate the double stranded DNA into single strands, break one, and pass the other through the gap to reduce or increase the linking number before repairing the broken backbone [24]. The DNA is therefore left with a different supercoiling density without permanent damage to the DNA backbone. Topo IIs, by contrast, break both strands of a dsDNA molecule, creating a gap which other dsDNA passes through. Depending on what DNA is passed through the broken strand, the supercoiling can reduce by 2, a positive supercoil can become negative, or two dsDNA duplexes can be interlinked [23]. This relaxation of supercoils has been shown to have important physiological functions at almost all stages of cell life, including packaging chromosomes, assisting with repair, replication, and transcription of DNA [25], making topoisomerases such as DNA gyrase in bacteria an attractive target for antibiotic drugs [26] and the type II topoisomerase in humans Top2 a target for chemotherapy [27].

Similarly central to all forms of life is the class of proteins known as the helicases, which unwind, unzip, and in some cases rewind DNA during replication or packaging [28]. They also can remove DNA bound proteins which may otherwise serve as a roadblock to other molecular machines [29]. Helicases differ from topoisomerases in that they do not break strands of DNA (and indeed often need broken covalent bonds (nicks) in the phosphate backbone to bind to [29]), but instead simply form single strands which may then be passed to or through other molecular machines.

Other molecular machines exist purely to repair errors in DNA. Whether that is repairing single- or double-strand nicks in the backbone or repairing DNA base excision, molecular machines exist which bind to the DNA and scan quickly along it looking for imperfections which need repair. These molecular machines also necessarily exert forces, usually on the piconewton scale, and lead to conformational shifts. Though these shifts may be small, or may affect only the torsional number of supercoils in the DNA, the effect on the cell cannot be overestimated. Each molecular machine is key to the cell, which is evident in their use as an antibiotic target [26] and the great number of illnesses which can be induced through the inhibition of molecular machines [25, 29].

These molecular machines work by transducing force and torque onto dsDNA, and the response of the molecule to this manipulation is a key part of their function. The mechanical properties of dsDNA are therefore biologically relevant parameters and understanding DNA response to applied force and torque is the first step to understanding the physics of these molecular machine/DNA interactions.

## 1.5 Simulations of DNA

Many schema have been devised to simulate DNA and other key biomolecules on different length scales, and modern hardware such as large RAM chips and highly parallelised general purpose graphical processing units (GPGPUs) allows simulations as demanding as quantum mechanical calculations of proteins making use of linear-scaling density functional theory [30]. At the other end of the scale, parameterisation and grouping atoms into units represented by one pseudoparticle reduces the complexity of a DNA model while keeping key features intact [31], making possible simulations of complex dynamical systems such as coarse-grained DNA walkers [32] which show the thermodynamics of a molecular machine composed of biased random motion.

It is useful in the context of this thesis to broadly characterise biological simulation as being of four types: quantum mechanical, atomistic, coarse-grained, or statstical. The first of these deals with length- and time-scales too small to be applicable in most experimental projects, and will not be discussed. However, atomistic and coarse-grained models are able to make testable predictions, and statistical or polymer physics models such as the wormlike chain [33] may be used to fit to experimental data to validate an instrument's measurement capability.

### 1.5.1 Atomistic Molecular Dynamics

Atomistic molecular dynamics concerns itself with constructing structures atom by atom and allowing them to interact through specified forcefields while maintaining the system at a given temperature and/or pressure. The mathematics of this process will be more fully explored in Chapter 4, but in general the procedure is as follows:

1. For each atom $i$ at time $t$ with velocity $\vec{v}_{i,t}$ and position $\vec{x}_{i,t}$, find and sum the forces due to every other atom according to the functional forms in the forcefield to give $\vec{F}_{i,t}$
2. Find the acceleration $\vec{a}_{i,t}$ due to this force using the atomic mass $m_i$: $\vec{a}_{i,t} = \frac{\vec{F}_{i,t}}{m_i}$
3. Numerically integrate $\vec{a}_{i,t}$ over the timestep $\delta t$ to give the new velocity $\vec{v}_{i,t+\delta t}$. Many schemes exist to do this, but with the Euler method for example $\vec{v}_{i,t+\delta t} = \vec{v}_{i,t} + \vec{a}_{i,t}\delta t$, $\vec{x}_{i,t+\delta t} = \vec{x_{i,t}} + \vec{v}_{i,t}\delta t$
4. Repeat from step 1

This idea is deceptively simple. In practice, forcefields are extremely complex, and rapidly become computationally expensive. Only recently, with specialised hardware [34] or advanced mathematical techniques [35] has it become possible to understand some systems' behaviour over millisecond timescales.

DNA interactions with other molecules, including solvents and ions, have been of interest to the biological simulation community since the field began in earnest and the solvation shells of DNA were elucidated [36, 37], showing that water interacted with DNA in an ordered, predictable way close to the DNA and inside the grooves. As the techniques have become more refined and hardware has improved, simulations have been run of DNA interacting with small biomolecules such as polyamines [38] demonstrating conformational changes as a result of polyamine binding. Other ligands such as transcription factors, which are proteins that bind to specific sequences of DNA to regulate gene expression, have been simulated over both short (150 ps) [39] and long (0.5-1 $\mu$s) [40] timescales. Though difficult and demanding, larger proteins' interactions with DNA fragments have become more and more possible and the state of the art in a review from a decade ago [41] are now routinely surpassed [42].

Both non-traditional (for example Monte Carlo [43]) and traditional [44, 45] means are tackling problems as fundamental as DNA-histone (large, alkaline proteins around which DNA is wrapped) mechanics [46]. Of perennial interest is the effect of molecular machines, but studying a complete molecular machine's interactions with dsDNA in atomic detail is still beyond modern computing.

Fortunately, elucidation of DNA conformational change in response to force and torque is not only possible through direct interaction with biomolecules. Instead, a small fragment of DNA may be readily simulated, and external forces on the molecule may be applied. This can be done with harmonic traps which apply a quadratic potential to a chosen particle centred on a chosen locus, positional restraints which do not allow any movement of a particle but which set all forces acting on that particle to 0, or more complex methods like boundary conditions (as done on DNA in a nanopore [47]) or steering of the simulation where a starting conformation is drawn towards a predefined end conformation by harmonic potentials acting on the atoms (as in [46]). Taken together, these make molecular dynamics a powerful tool for simulating the structure of DNA in stretching or twisting experiments, and indeed through careful simulation design it may in future be possible to replicate the conditions imposed on DNA by external agents in order to study them by proxy.

Stretching DNA with atomistic molecular dynamics has been of interest for as long as meaningful DNA duplexes could be simulated. Initial studies looked at stretching DNA to increase the distance between base pairs and so generate an intercalation site [48], but as time went on the focus turned more to elucidating structural transitions seen in experiment, and predicting novel DNA structures. Dual studies of DNA under tensile loads were performed [49] and a good numerical agreement between the extension lengths and forces of melting across methods lent credence and significance to the structures found in simulation. Novel structures were also studied in their own right, for example DNA induced to melt by application of force to the opposing 3' ends of the strands was seen to form a "ladder-like" structure under certain conditions now known as S-DNA [50], and again correspondence between the calculated and experimentally found force-extension behaviour indicated that this structure was plausible. Care must be taken, however, as with small fragments of DNA ($<30$ bp) the forces may not be reliable [51]. Further studies into the physics of the stretched DNA demonstrated that when calculating parameters of perturbed DNA the entropy is as important as other thermodynamic properties [51], and suggested entropic effects may account for the discrepancies seen between previous short-fragment simulations and experiments on DNA of several thousand base pairs.

Later investigations of torsionally unconstrained DNA under steadily increasing force [1] found a further structure in which the DNA underwound, lost its Watson-Crick hydrogen bonds, and formed a stable structure of bases stacked with bases in the opposite strand, maintaining stability by interlinking the two strands and forming a zip-like structure. Distinct force regimes DNA experiences while undergoing overstretching were also found.

Later dual simulations of DNA and RNA duplexes under load identified structural transitions and different force regimes, and demonstrated that it was structural differences between DNA and RNA which gave rise to different mechanical properties, specifically how the interstrand distance evolved through stretching was linked to different flexibility [52].

Overall, then, the realm of DNA stretching beyond its contour length (overstretching) has been reasonably well served by previous investigations and many new structures have been identified along stretching pathways. However, it is important to note that all studies of DNA under force have allowed the DNA to retain torsional freedom – it may rotate as much as it pleases to reduce its electrostatic potential energy. However, this is not necessarily always the physiological situation. On the contrary, for a molecular machine binding in the middle of a length of DNA, the DNA is most certainly torsionally constrained – or at least affected torsionally by the rest of the DNA. This omission is particularly glaring when compared to extensive experimental studies of torsionally constrained DNA, which are discussed in a later section.

The effect of torque on linear DNA structure has been relatively little studied, however. By imposing an angular restraint on the relative orientations of the first and last base pairs, the effect of added twist on the twist flexibility was studied [53], and the effect of base-pair sequence was well characterised showing that flexibility was ordered: pyramidine/purine steps were the most flexible, followed by purine/purine and purine/pyramidine. As in [1] the stacking interactions played a key part in dsDNA's mechanical properties.

Elsewhere, by making use of modified periodic boundary conditions, an effectively infinite tract of DNA was simulated, and over- and under-twisting achieved through deleting or adding base pairs. This metholodogy did not allow the DNA to form supercoils and was thus described as being in the "absence of writhe" – the DNA structural property describing how twisted around itself the DNA is. The distribution of the applied twist was analysed, and it was demonstrated that B-DNA partitions its perturbations such that there are regions of extreme structural disruption in coexistance with regions with a structure relatively close to the canonical B-form DNA structure, and that this process is sequence-dependent [12].

In many other computational investigations, rather than applying an external torque to the molecule with restraints or traps, the focus has instead been on generating torque internally to the DNA through over- and under-twisting closed circular DNA (ccDNA), a covalently bonded circle of DNA which is therefore internally torsionally constrained. To do this, the DNA twist parameter (i.e. the angular displacement between successive base pairs) when creating the ccDNA is varied such that the DNA undergoes fewer or more full helical twists over the total length of the minicircle [54]. Small DNA minicircles of 60-110 bp have been shown to bubble, kink, wrinkle, and supercoil as a result of applied torsional and bending stress [54]. Upon over- and under-twisting, larger ccDNA undergoes significant structural change and supercoils in a way dependent determined by the linking number $L_k$, which is the difference between the number of turns in the relaxed structure and that in the torsionally stressed structure [55], a relation that is shown experimentally to be a meaningful parameter [56]. This mechanical response, driven by the thermodynamics of the system partitioning energy between twist and writhe [57], brings close together sites which would normally be distant, impacting gene regulation [58], and mechanical response to torque may play a role in long-distance information communication [59]. Some initial work has investigated the effect of proteins and stiffer triplex regions on sites which are brought close together by supercoiling [60], which may again have impact on gene regulation. These findings elucidate the mechanics of torsion in DNA which is also under bending stress, as in a plectoneme. However, the effect of torsion in linear DNA is relatively unknown. The interplay between imposed torsion and stretching, which has been experimentally investigated, has yet to be interrogated *in silico*.

### 1.5.2 Coarse-Grained Molecular Dynamics

Rather than treating each atom individually, coarse-grained molecular dynamics seeks to replicate the physics and dynamics of a system with as few component parts as possible, allowing access to greater time- and length-scales than is possible in more detailed simulations. In general, a coarse-grained regime will identify regions of a system which can be modelled together as a single unit. For example, an amino acid in a protein may be represented simply as a point surrounded by a spherical excluded volume. One of the most successful coarse-graining methods for DNA is oxDNA [31], developed in the latter part of the last decade in Oxford. The system was designed to mimic as closely as possible the thermodynamic properties of DNA, and reduces each nucleotide to just two fictitious paricles – one representing the backbone/sugar ring and the other the base itself. Inter-bead forces were parameterised such that melting temperatures and similar parameters were in agreement with experiment [31], but from this simplicity interesting and complex systems like a DNA walker were well simulated [32], showing oxDNA can simulate complex, long-timescale thermodynamic processes.

Structural transitions are of particular interest in oxDNA simulations, both because it is important to be sure with such a coarse model that the dynamics mimic reality, and because the larger simulations allow insight into more experimentally relevant tracts of DNA. Duplex hybridisation has been well recreated [61], as has its opposite [62], and studies of significant bending [63] including DNA minicircles [64] reproduce well the atomistic insights from atomistic simulations and suggest new stable DNA conformations in the highly curved regions of plectonemes [65]. Overstretching of the DNA duplex has been studied with good agreement with other methods [66], though the authors point out the model does not recreate the expected ladder-like S-DNA structure characterised by its untwisted backbone and highly angled but intact base pairs, likely because there is physics not simulated which is responsible.

Over time the oxDNA model has become more sophisticated. Initially it was developed using only "average" bases , so no sequence dependence could be examined [31], however parameterisations were subsequently released which allow for the sequence to be taken into account [67]. Similarly, the major and minor grooves of the double helix were not initially represented, but new developments [68] allow its effect on DNA in this model to be studied [69].

Salt concentration is one of the most physiologically important parameters because it has been shown to have dramatic effects on the flexibility of DNA [70] as it electrostatically screens the system. Initially only high salt conditions were simulatable due to hard-coded values in the solvation system, but support for user-defined salt concentration was added in the same forcefield as major and minor grooves [68].

Similarly to the atomistic realm, however, relatively little has been done looking exclusively at twisting DNA. One such study looked at the thermodynamic properties of twisted oxDNA DNA [71], and built a model of the free energy of local defect production due to torsional stress. By modelling the free energy as a local interaction between neighbour base pairs plus a long-range correction, the positions of defects formed by introduced torsional stress was predicted. Another recent paper demonstrates that in agreement with polymer physics models oxDNA produces a wave of twist under certain conditions [72] in which regions of high twist propagate through the molecules. Despite this, and the extensive work on DNA machines and conformation done with oxDNA, a complete picture of DNA mechanics and structure under only twisting is yet to emerge. Similarly, the effect of stretching while twisting DNA remains under-explored, possibly because the simulations themselves are difficult to design, require multiple individual simulations, and the complex interactions produced by such perturbations resist straightforward analysis.

### 1.5.3 Polymer Physics and Statistical Mechanics

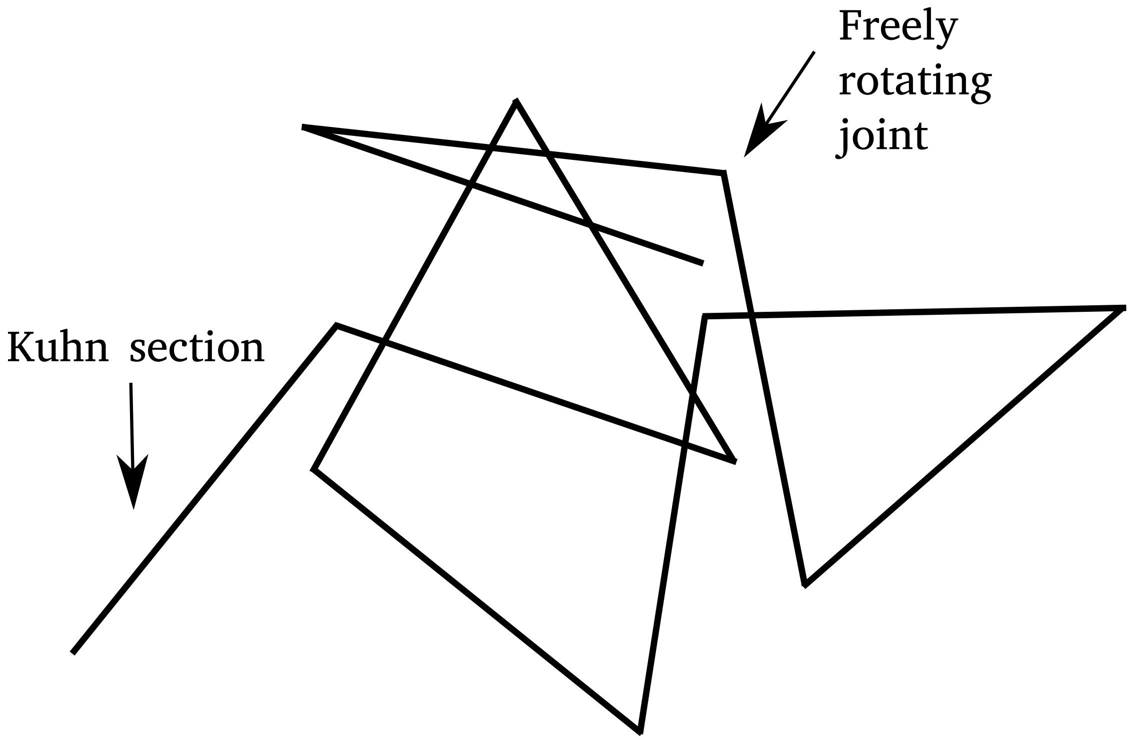


Figure 1.4: A Cartesian representation of a 2D freely-jointed chain polymer consisting of 10 Kuhn sections

The behaviour of DNA is inextricably linked to the length scale of the molecule under consideration, and as seen in previous molecular dynamics studies care must be taken with short fragments that the length of the DNA is sufficient to demonstrate the desired behaviour. At the other end of the scale, DNA which is thousands of base pairs long may be well described by various statistical mechanics or polymer physics frameworks. One of the simplest of these is the freely jointed chain model which models the DNA as being made of individual links. While the links themselves are inflexible, they are free to rotate around the joints between them, and these rotations are completely independent of each other. The DNA conformation is therefore analogous to a random walk, changing direction after each “step” or inflexible segment. This allows complex conformations to appear and causes the DNA to behave as an entropic spring as seen experimentally – as the ends of the polymer are stretched apart, there are fewe microstates for the DNA to occupy. As systems act to maximise their entropy, there is a force which resists the stretching. Differing stiffnesses of polymers is modelled by changing the length of each link, also known as the Kuhn length. The shorter the Kuhn length, the more flexible the polymer. An example of a freely-jointed chain polymer is shown in Figure 1.4.

The freely jointed chain is an extremely simple model and therefore misses some physics present in a real DNA system, however. A more complex model which can take account of twisting and stretching beyond the contour length, known as the wormlike chain (WLC), is commonly used to fit to experimental data and develop theoretical descriptions and predictions of DNA behvaiour. In this description, the polymer is imagined as a continuous flexible rod whose flexibility is governed by the persistence length. The simplest WLC model was introduced in 1995 [73] and had the form

$$F = \frac{k_B T}{L_P} \left\{ 0.25 \left( \frac{x}{L_C} \right)^{-2} - 0.25 + \frac{x}{L_C} \right\} \tag{1.1}$$

where $F$ is the end-to-end force on the DNA, $x$ is extension, $k_B$ is the Boltzmann constant, $T$ is temperature, and $L_P$ and $L_C$ are persistence and contour lengths, respectively.

Modifications have been made to this basic functional form both to improve the accuracy of the force-extension fit [74], and to allow the model to include enthalpic and entropic terms in low- [73] and high-force [74, 75] regimes. Further, the WLC polymers can be twisted and the stretch-twist relation accurately characterised [76] or the twist stiffness calculated [33]. Due to the simplicity and robustness of the model, fitting the WLC and comparing calculated persistence length with other experiments has become a standard tool in force transduction experiments to verify that only one molecule is being manipulated, or to assess the change in persistence and contour length due to ligand binding or other experimental conditions.

## 1.6 Experimental Approaches to Measuring and Applying Forces

Various methods have been developed to apply piconewton scale forces to molecules and cells *in vitro*. In this section the main work will be summarised and the methods compared.

### 1.6.1 Magnetic and Optical Tweezers

Optical tweezer experiments make use of the fact that a bead with optical density higher than its surrounding medium will be attracted to the centre of a focused laser beam, and in the centre of the trap there is a region in which the force applied scales linearly with bead displacement in the manner of a simple spring. Most commonly, the beam profile used is Gaussian but it is possible to use other beam profiles for specific applications. The physics of optical trapping is given in Chapter 2.

Briefly, the general principle of optical tweezing is to attach the molecule of interest to at least one optically transparent dielectric particle, commonly polystyrene or latex. The molecule may be free at the other end, attached to a cover slip surface, or attached to a second optically trapped bead in a configuration known as a dumbbell experiment [77]. The molecule then has force applied to it either by applying flow or electric fields in the case of the singly attached molecule experiments, or by moving one or both traps for the experiments with both ends of the DNA fixed. The forces on the molecule may be measured through a suite of methods, the most common of which is back focal plane detection [78]. The trapping laser (or a separate tracking laser) is focussed on a quadrant photodiode (QPD), which produces a voltage when exposed to light. The voltage difference between the two upper quadrants and two lower quadrants, and two leftmost quadrants and two right quadrants, give a measure of the position of the bead in the trap by indicating where its shadow is cast. From Hooke's law, with a known trap stiffness, the force applied may then be calculated. The trap position may then be algorithmically adjusted to maintain a desired constant force (a force clamp) [79], or moved to vary the force applied. Alternatively, the trap may be fixed in position and used to measure the forces produced by DNA interactions with other biolomecules; they may also be left fixed to allow imaging of biomolecular interactions with fluorescence microscopy.

Magnetic tweezer experiments work in a similar way. A bead embedded with a magnetic metal like iron oxide is attached to a tether which is attached at the other end to a surface. By application of a nonuniform magnetic field (i.e. $\vec{\nabla}\vec{B} \neq 0$) a force may be applied, by application of a rotating field of uniform strength ($\vec{\nabla}\vec{B} = 0$) the direction of the magnetic dipole of the bead may be manipulated and thus a torque generated.

#### 1.6.1.1 Pulling on and Twisting Tethers

The first experiments were performed primarily in a brightfield configuration – that is to say, no fluorescence was used and so the DNA itself was not imaged. Instead, its behaviour inferred from force, torque, and extension measurements. This proved to be very successful, and a great deal of information was extracted, including elucidation of structural transitions..

The first mechanical measurements of single molecules of DNA were performed in

1992 [80]. DNA was tethered between a magnetic bead at one end and the cover slip surface at the other. By applying a constant flow to stretch the DNA molecule out, and a magnetic field with a permanent magnet, forces up to 20 pN were applied, and brightfield microscopy used to track the position of the bead and hence produce force-extension curves. These were broadly in agreement with the FJC model at forces lower than 1 pN but quickly diverged, a disagreement which subsequently formed the basis of the WLC model. The authors also gathered fluorescence microscopy images of the DNA tethers.

More force-extension work quickly followed, with the properties of DNA examined in greater detail than had been possible previously [81]. Optical traps were also used to investigate DNA unzipping under force and the sequence dependence thereupon [82], which showed that unzipping CG pairs led to a spike in force applied compared to unzipping AT pairs, explained by GC pairs having three hydrogen bonds to AT's two. DNA's interactions with the dye YOYO-1 and the protein RecA were also examined through tethering between a micropipette and an optically trapped bead [83], with the change in contour length upon binding investigated. They concluded that RecA binding led to a 45% increase in DNA contour length, and added YOYO-1 increased the contour length by a further 10-20%. Other work on YOYO-1/DNA complexes showed increases in contour length due to YOYO-1 alone of between 28 and 38% [84, 85, 86].

Tethering to the surface, however, is not the most stable trap configuration as mechanical noise can be introduced through the stage and slide holder. A powerful means of removing that possibility is to fix the DNA between two optically trapped beads in solution. In this configuration, the thermal fluctuations of DNA end-to-end force was measured to unprecedented femtonewton precision [77] (some three orders of magnitude below typical physiological forces), DNA was tied in knots [87], and the activity of RNA polymerase moving along a DNA template was observed to near base pair resolution [88, 89].

Similar studies on DNA dumbbells showed the binding affinities of ligands (for example [90]), and an ambitious experiment demonstrated the possibility of studying protein trajectories along DNA tethers by combining STED and confocal microscopy and a DNA dumbbell configuration [91]. The effect of tension on unpeeling and unwinding of DNA has also been quantified by experiments using a DNA dumbbell with one strand nicked at one end. The unpeeling of the nicked strand was dramatically visible through fluorescence microscopy and the resulting untwisting was modelled using the twistable wormlike chain [76].

Many experiments have also been carried out to study either the effect of added twist to DNA systems or the effect of added ligands or molecular machines to the DNA twist. One early experiment demonstrated how the properties of DNA stretching and the extension of the DNA molecule change under a known tension with different degrees of supercoiling [92]. In this seminal work, it was shown that for all applied forces, overtwisting the DNA results in a shortening of the tether after a force-dependent plateau. However, for forces above 0.8 pN, undertwisting the DNA had no effect on the overall length. The number of added turns needed to shorten the tethers, and the subsequent rate of shortening, by either undertwisting or overtwisting was shown to be dependent on the tension applied to the molecules, with a higher force resulting in increased overtwisting needed before significant shortening was seen. In similar work, the transition between the slowly-shortening low-twist regime and the rapidly-shortening high-twist regime was studied [5]. It was shown that the reduction in length due to applied rotation can be modelled as two linear regions separated by an sudden transition known as the buckling transition. The appearance and severity of these buckling transitions was shown to be a function of applied force, and may be a good metric to use for broader torsional properties.

In others, different structural transitions of overwound or torsionally constrained DNA under stretching were predicted [93], overwound and stretched DNA was shown to assume a structure if supercoiled and stretched in which the backbones turned inwards and the bases themselves face out [2]. Molecular machines were studied and the action of a type II topoisomerase relaxing applied supercoiling was examined [94], as was the overtwisting and bending of DNA due to topoisomerase IV [95], a bacterial type II topoisomerase similar to DNA gyrase.

Monitoring the DNA twist as opposed to applying supercoiling led to the counterintuitive discovery that upon stretching DNA overtwists [96], and various schema have been devised to measure the twist fluctuations of DNA such as [97, 98], which monitored the position of a gold nanoparticle and the motion of a tethered bead which was free only to rotate, respectively. The gold nanoparticle was used as a direct measurement of twist, while the freely orbiting bead was tracked and its trajectory used to infer the twist on the DNA. As theoretical studies remind us [99], however, it is important to be certain that a genuine phase transition is being observed – looking at only individual parameters like extension was shown to give incorrect onsets of phase transitions. Magnetic tweezers have also been used to examine the effect of binders such as ethidium bromide and cisplatin on the mechanical properties of DNA [100], which showed that upon binding the DNA experienced a ligand-specific shift in its contour length, bending, and inter-base pair twist.

DNA/RecA fibres under tension have been found with magnetic tweezers to have torsional stiffness which is highly force-dependent [101]. The mechanical properties of DNA under twisting have been shown to be extremely dependent on salt concentration and type, including the twist rigidity [102] in which a cohesive prediction of DNA twisting rigidity was not reported such were the variations. The buckling torque as a function of sodium chloride concentration has also been investigated [103], and it was shown that the buckling torque reduces with increased salt, presumably due to higher electrostatic screening and shorter persistence length.

Finally, recent work has demonstrated the dynamics of plectonemes. By forming plectonemes with a standard vertical magnetic tweezer and then moving the magnet such that the tether lay horizontally to the field of view, plectonemes were shown in fluorescence to be highly mobile and move along the DNA tether with speeds of micrometers per second. However, due to making the plectonemes in a vertical orientation, the actual process of plectoneme formation could not be observed, though it was deduced that the location of plectonemes depended on the DNA sequence.

Overall, tether based experiments have given great insight into the dynamics and mechanical properties of not only DNA but DNA-ligand systems. However, combining both force and torque measurements of DNA-protein dynamics with real-time fluorescence microscopy has so far proved elusive.

#### 1.6.1.2 Flow- and Electric Field-Based Experiments

Human ingenuity is a remarkable thing, and many other methods to apply forces to DNA beyond tweezing have been devised. Here some examples will be briefly introduced and their limitations identified.

One of the early experiments imaging DNA held with optical tweezers was performed by Perkins *et al* and published in 1994 [104]. Single molecules of DNA of varying lengths were attached to 1µm diameter beads which were trapped optically. The trap was then moved to induce flow around the bead and hence extend the DNA. The DNA was fluorescently imaged and the relaxation rates measured and found to be in good agreement with previous studies. This transformative work did however have some drawbacks, namely that flow was necessary to image the DNA, and that no external force other than the drag could be applied to the molecule. For directly studying protein-DNA interactions or DNA dynamics under forces, this method is sadly lacking.

Similarly, DNA can be stretched out by a meniscus of fluid moving past it. In 1995 this met with some success, with the tensile strength of the DNA estimated to be 476 pN [105]. However, the efficacy of the technique was surface-dependent and the surface-attachment of the DNA means that once again this is not a suitable technique for monitoring DNA dynamics through time.

DNA is an electrically charged molecule and is therefore susceptible to electric fields. Using an ingeniously designed artificial nanopore, very small segments (3-10 nm) of a DNA construct have been stretched [106]. Though that stretching was well modelled with molecular dynamics simulations, the lack of ability to image the DNA directly limit the scope of the investigations which can be carried out, and it is unsuitable for a combined approach as in this thesis.

The tip of an atomic force microscope cantilever can be attached to a single DNA molecule in a similar way to a micropipette. The cantilever can then apply and measure force and provide insight into stretching and phase transitions. However, the geometry of a typical AFM DNA experiment is in a vertical configuration – the DNA held between the cover slip and the cantilever directly above it. This makes direct imaging of the DNA an extremely hard task, though possible [107]. However, there is no ability in these experiments to apply torque and therefore is not a suitable method to use for combined force and twist transduction.

Finally, the fictional centrifugal force has applications beyond astronaut training. By putting DNA on a spinning disk, it is possible to stretch out the molecule and attach it to the surface. This can then be fluorescently imaged or data recorded with atomic force microscopy [108]. Despite this approach's success in imaging surface immobilised DNA, the lack of dynamics make it unsuitable for seeing DNA processes as they occur.

### 1.6.2 Comparison of Methods

Magnetic and optical tweezers each have many strengths and some weaknesses. Optical traps are versatile and can apply hundreds of piconewtons of force to a DNA tether, but cannot easily apply torque (birefringent particles and polarised light do make it possible, however [99]); further, in the case of a DNA dumbbell experiment, both ends of the DNA are rotationally unconstrained and so application of torque is impossible. However, optical traps are exceptionally good at generating tethers which are more or less horizontal to the field of view, and so imaging tethers is possible. In the case of the dual-trap experiment, the DNA is guaranteed to be held transverse and so imaging protein-DNA dynamics is made relatively conceptually straightforward.

By contrast, magnetic tweezers excel at introducing torsion to systems, but in general apply lower tensions than are accessible to optical traps. The configuration for a magnetic tweezer experiment is most usually vertical – that is to say the bead is directly above the DNA as seen in the microscope. This allows twist monitoring as seen above, but precludes straightforward imaging of the DNA itself.

Each method is powerful in its own right, and combining approaches to make use of the transverse nature of optically trapped tethers and the rotational properties of the magnetic trap is a gap in the current research, which, if filled, would enable extremely detailed studies of biomolecular interactions with DNA, RNA, and other filamentous molecules.

# Chapter 2

# Experimental Techniques

This chapter describes the theoretical foundations and practical considerations which go into designing and constructing magnetic and optical tweezers and a fluorescence microscope. It also details the methodologies used to gather data and extract meaning, and the specific method development including optimisation of the surface attachment of the 5 $\mu$m antiDIG beads, DNA conjugation to these immobilised beads, formation of bead-bead tethers, and labeling and imaging the DNA with YOYO-1.

The initial design, simulation, and construction of the magneto-optical tweezers was performed by Dr Zhaokun Zhou [109], then a PhD student in the Leake group, and subsequent refinements and rebuilds were performed by Dr Zhou and the author together. The fluorescence path was designed and constructed by Dr Helen Miller, then a PhD student in the Leake group, with subsequent limited and very minor modifications made by the author. The corrective module described to reposition the focal plane and therefore bring an optically trapped bead into focus is the work of the author. The preliminary experiments presented here which demonstrate correct optical trapping and magnetic rotation functionality were performed jointly with Dr Zhou. Software in Python 2.7 using SciPy, NumPy and Matplotlib was written by the author to obtain optical trap stiffness, and produce and display force-extension and force-twist curves. Dr Zhou wrote the LabVIEW software which performed data acquisition and force clamping.

The design and construction of the DNA handles and 4 kilobase-pair (kbp) DNA tether was undertaken prior to the commencement of the project by Dr Zhaokun Zhou based on work previously published by Yeonee Seol [110, 111], and the construct was then tested by the author using gel electrophoresis. The 15 kbp tether was design and constructed with Dr Jamieson Howard of the Department of Biology at the University of York and again verified by the author with gel electrophoresis. Therefore, the process of designing and making the 4kb DNA construct is described only briefly while the 15 kbp tether is described in more detail. The full protocol for constructing the 4 kbp tether is reproduced from Dr Zhou's PhD thesis in Appendix A. All other work in this Chapter is the author's.

## 2.1 Target Experimental Geometry

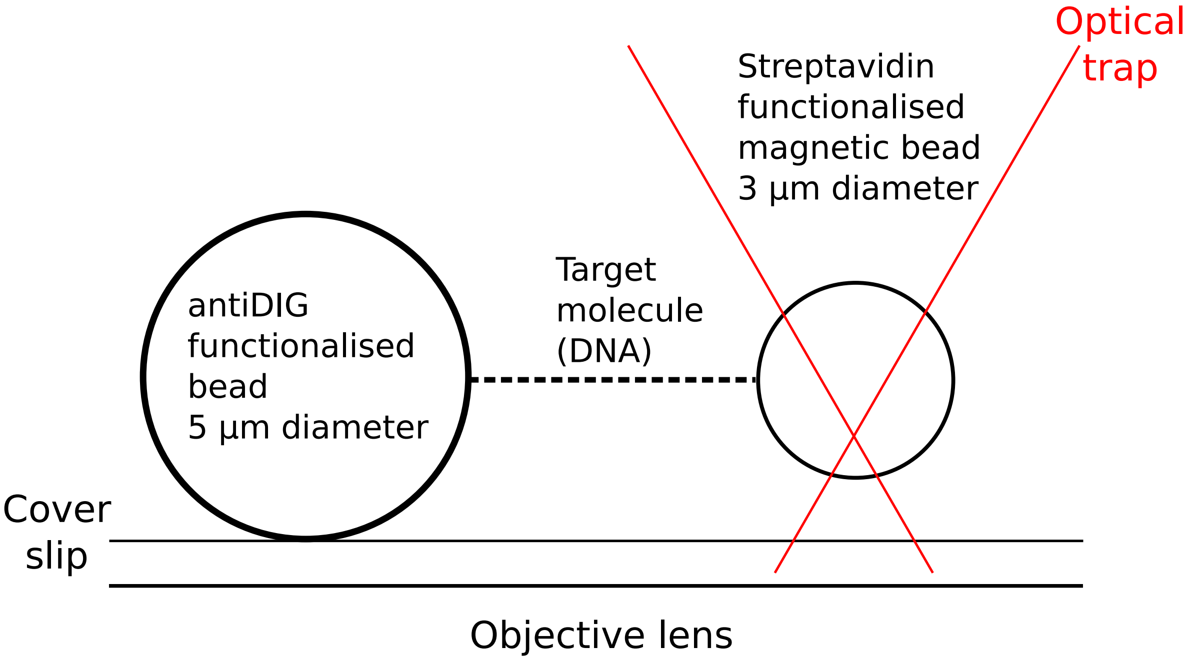


Figure 2.1: The target experiment. A 5 $\mu$m antiDIG functionalised bead is immobilised on the cover slip surface. A 3 $\mu$m streptavidin-coated bead is trapped optically. The target molecule of interest forms a horizontal tether between the two beads. By applying force with the optical trap or torque using bespoke Helmholtz coils mounted on the microscope (not pictured), the tethered molecule can be directly manipulated. As the tether is transverse to the objective lens, the full contour of the molecule is visible.

The aim of this experimental work was to design and implement a novel magneto-optical tweezing geometry to allow transverse tweezing and hence imaging of the full contour length of the studied molecule. Briefly, this was achieved by use of a 5 $\mu$m anti-digoxigenin (antiDIG) coated bead which was attached to the cover slip surface. To this was conjugated DNA constructs which were functionalised at one end with multiple digoxigenin molecules for attachment to the 5 $\mu$m antiDIG bead and at the other with biotin. During the experiment, a 3 $\mu$m streptavidin-coated magnetic bead was trapped in a near-infrared optical trap and brought to within 500 nm of the immobilised beads. The free end of the DNA then attached to the 3 $\mu$m bead through the biotin/streptavidin interaction, forming a tether. By moving the nanostage the two beads were then separated, leaving the tether horizontal to the field of view.

Magnetic control was effected by way of two pairs of Helmholtz coils mounted on the microscope body. Driving each with an alternating current led to a rotating magnetic field which induced rotation in the trapped magnetic bead and hence applied additional twist to the DNA tether.

The target experimental geometry, excluding the Helmholtz coils, is shown in Figure 2.1.

## 2.2 DNA Design

The structure of the two DNA constructs used in this project follow the same pattern. A stretch of dsDNA is attached at each end to a 500 base-pair (bp) handle functionalised either with biotin or digoxigenin for conjugation to streptavidin- or antidigoxigenin-functionalised beads, respectively. The biotin/streptavidin link is commonly used for its strength: it is the strongest non-covalent force found in nature [112]. Digoxigenin (DIG) is an antigen which binds extremely strongly to its antibody anti-digoxigenin (antiDIG), and because of its small size and relative ease of attaching to nucleotides it is a common choice. Two different constructs were produced for this project, a 4 kbp construct of approximately 1.3 $\mu$m contour length and a 15 kbp construct of approximately 5 $\mu$m. These lengths were chosen so that there should be sufficient length between the beads to image the DNA with fluorescence microscopy and to see the change in contour length due to under- and over-twisting.

### 2.2.1 The 4 kbp Construct

There are two methods to construct a DNA molecule of known sequence. Either a sequence of DNA can be cut out from a larger construct, or the DNA can be made from individual nucleotides using the polymerase chain reaction (PCR). PCR is a method whereby a plasmid – a circular DNA construct of known sequence – is used as a template for DNA replication. The replicated DNA then has the same sequence as the plasmid, and multiple cycles of this replication lead to high yields. In order to have a given length of DNA reproduced, it is necessary to design primers, which are short (∼20 base) fragments of DNA which anneal to the plasmid in a sequence-specific way. This is necessary because DNA polymerase is unable to begin from nothing – there must be DNA for it to bind to and then extend. The primers are designed in such a way that the polymerases move towards each other, replicating the region between the two.

As in the work of Yeonee Seol [110], the DNA handles were prepared from the pBluescriptII (+) plasmid. Primers were designed which bracket a 500 bp stretch of the plasmid (see Figure 2.2), and this sequence was amplified through PCR in the presence either of biotin-labeled dUTP (Biotin-16-dUTP, Jena Bioscience) or a digoxigenin-labeling dNTP mix (DIG DNA labeling mix, Roche) [111]. The concentrations of biotin-16-dUTP were calculated such that in the final biotin-labeled handle ("biotin handle") there are an average of 85 biotin-labeled nucleotides and on the digoxigenin-labeled handle (hereafter "DIG handle") an average of 30 digoxignenin-functionalised nucleotides [111]. This density of labeling should be sufficient to form strong bonds with the functionalised beads, and to have multiple binding sites on each strand and thus keeping the DNA torsionally constrained.

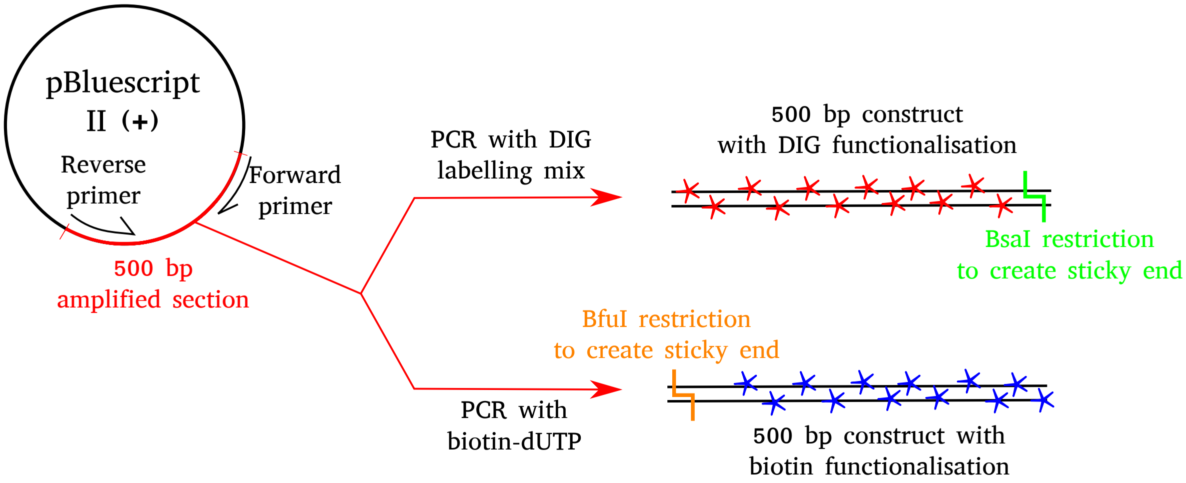


Figure 2.2: Schematic diagram of the PCR and restriction undertaken to create the 500 bp handles for the 4 kbp construct. A 500 bp stretch is amplified from the plasmid pBluescript II (+) in the presence of either DIG labelling mix or biotin-dUTP to create 500 bp constructs which will conjugate to antiDIG or streptavidin coated beads respectively. Each 500 bp construct is restricted to create a 4 nucleotide overhang which will join to a complementary overhang on one end of the 3 kbp midsection.

The handles as produced by PCR have blunt ends – that is to say, the strands terminate in the same place – and would therefore be difficult to ligate to the DNA midsection. In order to facilitate the final construction, four nucleotides were removed from one end of the DNA strands. This leaves a 4 nt overhang which will more easily attach to complementary sites on the DNA construct midsection. The biotin handle was restricted with the restriction endonuclease BsaI and the DIG handle with BfuAI, as in Figure 2.2. The endonucleases are a class of enzymes that cut the DNA backbone after binding to a specific base pair sequence. The cuts in the DNA allow the 4 nts at the end of the construct to escape, leaving the desired overhang.

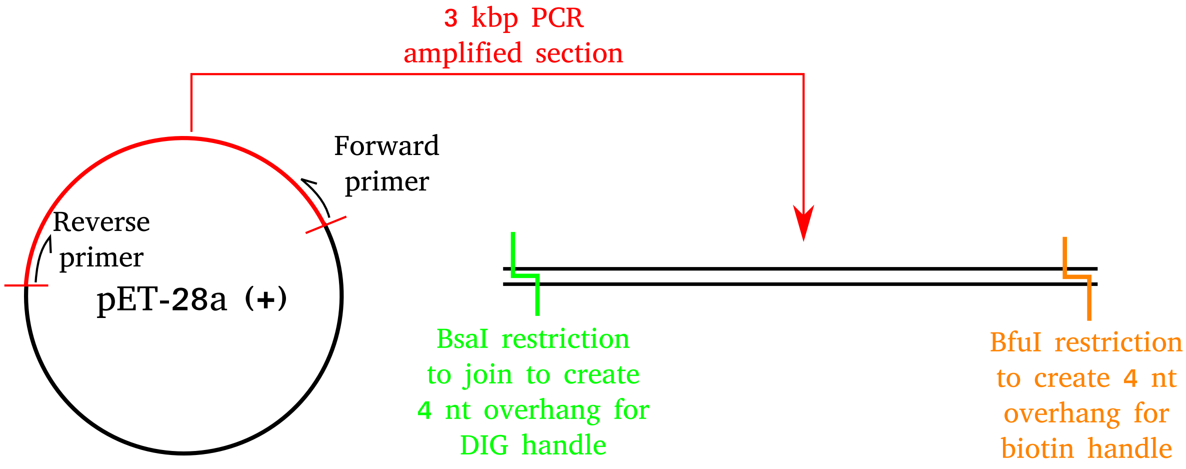


Figure 2.3: PCR schematic for the midsection of the 4 kbp construct. a 3 kbp section of the pET-28a(+) plasmid is amplified via PCR and restricted with both BfuI and BsaI to create a 4 nt overhang on each end. These overhangs are complementary to those created on the functionalised 500 bp handles so that one of each handle anneals to each midsection.

The midsection for the 4 kbp construct was based on the pET28a (+) plasmid, again as in [110]. Primers were designed to attach at either end of a 3 kbp tract and this section was amplified with PCR [111], depicted pictorially in Figure 2.3. On completion of the PCR, the constructed midsection was digested with both BsaI and BfuAI to create complementary nucleotide overhangs for attachment to the DNA handles. The joining of the handles to the DNA midsection was performed by mixing the handles in a 10:1 excess with the midsection and ligating with T4 DNA ligase (M0202S, New England BioLabs Inc.) [111]. This excess of handles ensures that all midsections will be attached to both handles, maximising the yield of the final product. The full protocol for construction of the 4 kbp construct may be found in Dr Zhou's PhD thesis [111] and is replicated in Appendix I.

Gel electrophoresis was then used to verify that the PCR had succeeded. Gel electrophoresis is a method whereby a DNA sample is loaded into a well at one end of a gel substrate which is immersed in an electrolytic buffer. The gel has holes of various dimensions, allowing different sized molecules to move through at different rates. A potential difference is put across the substrate, and as DNA is negatively charged, it moves towards the positive electrode. After some time, the DNA sample is imaged with a fluorescent dye such as ethidium bromide or SYBRSafe. The bright bands of DNA seen in each lane are then compared to a ladder – a control run alongside the sample which has bands at known molecular weights.

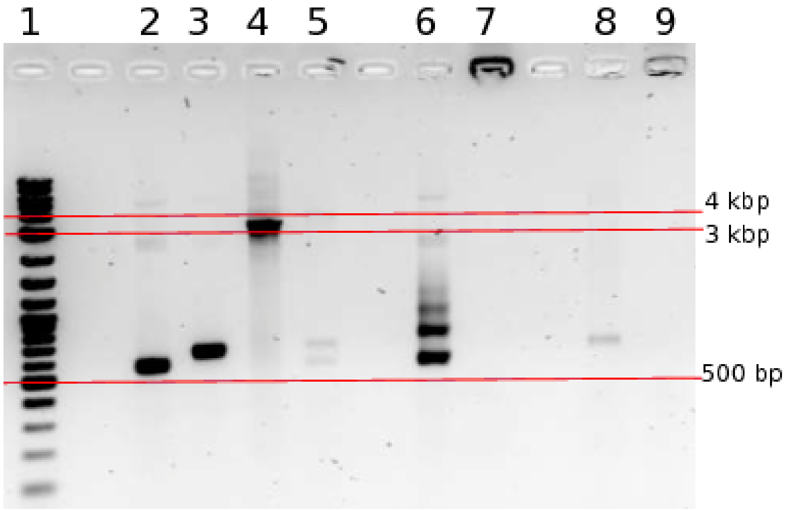


Figure 2.4: The 4 kbp construct tested with electrophoresis on a 1% agarose gel. Lanes: 1. 2-log ladder; 2. The DIG handle; 3. The biotin handle; 4. The 3 kbp midsection; 5. The full construct; 6. The DIG handle with antiDIG; 7. The biotin handle with avidin; 6. The full construct with antiDIG; 7. The full construct with avidin

Agarose gel electrophoresis was performed on each handle, the 3 kbp midsection, the final product, and the final product in the presence of molecular avidin and antidigoxigenin. The DNA was run on a 1% agarose gel at 100 V for 35 minutes in Tris-Acetate EDTA (TAE) buffer, measured against a commercial molecular weight ladder (N3200S, New England BioLabs Inc.) and imaged with SYBRSafe (S33102, ThermoFisher Scientific). As can be seen in Figure 2.4, in lanes 2 and 3 the DIG handle and biotin handle are seen as bright bands just above the 500 bp ladder band (at 600 and 700 bp weight respectively) confirming the success of the PCR, and in lane 4 the 3kbp midsection is brightly seen at the correct weight. In lane 5 the midsection band shifts upwards to around the 4 kbp band showing that the ligation has been successful, and the excess handles are seen at below it as expected. In lane 6 with the addition of molecular antiDIG the handles are shifted up demonstrating correct activity, while in the presence of avidin the handles form a large complex and remain in the well. In lanes 8 and 9, the tether is seen with molecular antiDIG and avidin respectively. In lane 8 the construct band has disappeared, and in lane 9 the large complex has again been formed and remains in the well, indicating that conjugation to the molecular antibodies was successful and the activity of the functionalised nucleotides is as expected. The 4 kbp DNA construct can therefore be used with confidence.

### 2.2.2 The 15 kbp Construct

Fabricating DNA tracts with PCR becomes less accurate as the length of the desired tract increases. For longer stretches, unless a specific base-pair sequence is desired, it can be a more robust method to take a premade longer DNA contour and digest it to the desired length with restriction enzymes. For the 14 kbp midsection required to make a 15 kbp construct, commercially available lambda DNA (N3011S, New England BioLabs Inc.) was used as a template.

The single cutters (restriction enzymes with exactly one binding site) of lambda DNA (a standard 48.5 kbp DNA construct derived from *E. coli*) NgoMIV (R0564S, New England BioLabs Inc.) and NheI-HF (R3131S, New England BioLabs Inc.) were used to digest the lambda DNA into three sections with lengths 20040 bp, 14639 bp, and 13823 bp. The 14.6 kbp middle section is the only one restricted by both NgoMIV and NheI-HF, and therefore had the desired 4 nt overhangs at each end. It was this section that went on to make the full construct. However, all three sections are very close in length, and so the two unwanted products could not be separated from the target sequence. In the final product, this meant that there were effectively three separate constructs:

1. The desired 15.6 kbp construct functionalised at both ends
2. An undesired 20.5 kbp construct functionalised at only one end
3. An undesired 14.3 kbp construct functionalised at only one end

While it would be better if possible to remove all constructs which are not the designed one, in practice it is not a cause for alarm for the unwanted DNA to be present. As they are functionalised only at one end, it should not be possible for the by-product constructs to form tethers between the two beads, and in an experiment their presence should therefore not be noticed.

To make the three fragments, 25 $\mu$g of the lambda DNA was restricted by 40 units of NheI-HF and 20 units of NgoMIV, in a total volume of 100 $\mu$L. CutSmart buffer (B7204S, New England BioLabs Inc.) was used as both enzymes are active in it. The digestion was performed for 1.5 hours at 37°C.

On completion of the restriction reaction, the enzymes were removed using ethanol precipitation. 10 $\mu$L 3M sodium acetate at pH 5.2 was added, along with 300 $\mu$L 70% ice-cold ethanol, and the resulting solution was stored at -80°C for between 1 and 2 hours before being centrifuged at 4°C in a bench-top centrifuge at maximum speed for 10 minutes. The supernatant was then discarded, and the remaining pellet suspended again in ice-cold 70% ethanol. Again the mixture was centrifuged at 4℃ for 10 minutes and the supernatant discarded. The pellet was then resuspended in 10 mM Tris-HCl buffer at pH 8.5 to a volume of 100 $\mu$L.

The handles were prepared similarly to those in sub-Section 2.2.1. However, as NheI-HF and NgoMIV were used for the restriction of lambda DNA, they also needed to be used to create the handle overhangs, necessitating redesigned primers. The primers were again designed to encompass a 500 bp section of pBluescriptII (+). The forward primer was designed for use with NgoMIV and therefore had to include its restriction site GCCGGC, while the reverse primer was designed for use with NheI-HF and included the sequence GCTAGC. The final primers had sequence

- Forward: $5'$ AAT TCG CCG GCT GGC GTA ACA TGG TCA TAG $3'$
- Reverse: $5'$ AAT TCG CTA GCT TTA TAG ACC TGT CGG GTT TC $3'$

The PCR reaction was conducted in Phusion HF buffer (B0518S, New England BioLabs Inc.), with 0.5 $\mu$M of each primer, 200 $\mu$M nucleotides, 200 ng pBluescriptII (+), 3% dimethyl sulfoxide (DMSO), and 10 units Phusion polymerase (M0530S, New England BioLabs Inc.) in a total volume of 500 $\mu$L. Two PCR reactions were performed with different nucleotides present to form the differently functionalised handles. For the biotin handle, Biotin-16-dUTP was again used, with the nucleotide mixture consisting of 200 mM dATP, 200 mM dCTP, 200 mM dGTP and 100 mM each dTTP and biotin-16-dUTP. This resulted in the final handle having approximately 120 biotin-functionalised nucleotides. For the DIG handle, the nucleotide mix again consisted of 200 mM dATP, 200 mM dCTP, 200 mM dGTP, but this time had 27 mM digoxigenin-11-dUTP and 173 mM dTTP, for a final labeling density of approximately 30 digoxigenin-11-dTTP per 500 bp handle.

The 500 $\mu$L mixtures were then divided into 20 tubes of 25$\mu$L to facilitate quick heating and cooling during the PCR cycle. The PCR reaction itself was performed according to the following procedure:

1. Heat to 98℃ and hold temperature for 30 seconds
2. Hold PCR mixture at 98℃ for 10 seconds to melt double stranded DNA, leaving single strands for the primers to anneal to
3. Bring PCR mixture to 60℃ for 30 seconds to allow annealing of the primers to the single-stranded template DNA

4. Bring mixture to 72°C for 30 seconds for extension. The standard approximation is 1 minute/kbp, so for a 500 bp construct 30 s was used.
5. Hold mixture at 72°C for 5 minutes for the final PCR extension to take place
6. Cool to room temperature to finally anneal single-stranded constructs into dsDNA

Steps 2-4 inclusive (in blue) were repeated 35 times to maximise product.

On conclusion of PCR, the handles were purified using the QIAquick PCR Purification Kit (28104, QIAGEN) according to manufacturer instructions. To form the overhanging nucleotides for construct ligation, the biotin handle was then digested using NheI-HF and the DIG handle with NgoMIV following the same procedure as that for the lambda DNA above. Once again, the products were then purified using the QIAquick PCR Purification Kit.

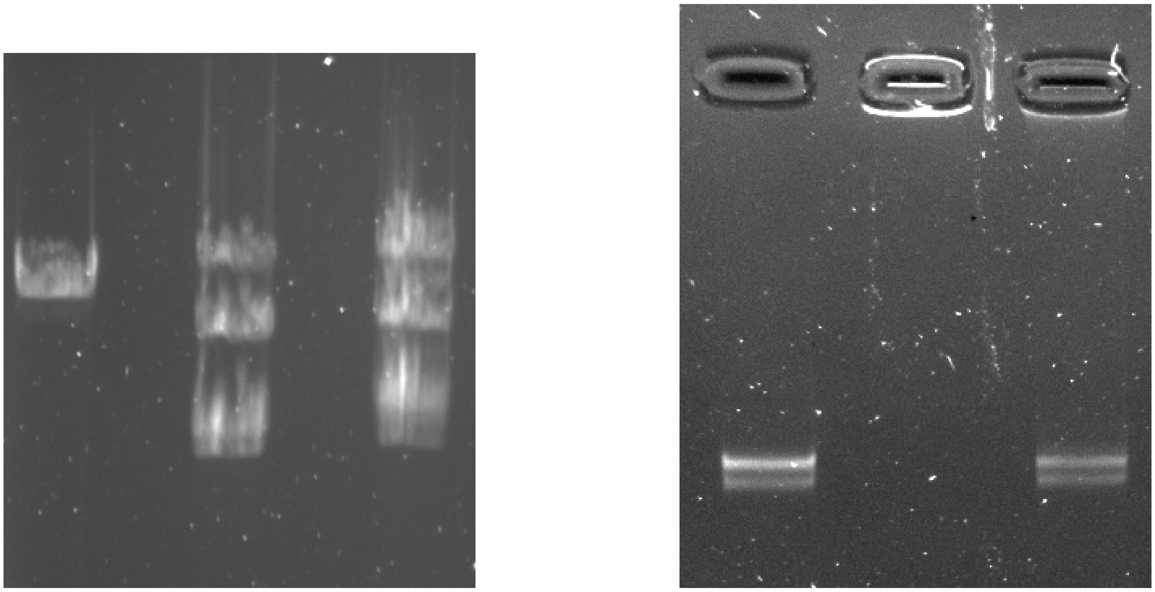
Similarly to the 4 kbp construct, a fivefold excess of handles was used for the annealing and ligation of the full construct in order to ensure that all midsections were fully functionalised. To anneal the products together, the midsection and handles in a ratio of 1:5 were heated to 65°C for 10 minutes and allowed to cool slowly to room temperature using a programmable heatblock. The backbones were then ligated with 1000 units T4 DNA ligase (M0202S, New England BioLabs Inc.) at room temperate for 10 minutes in T4 DNA ligase buffer (B0202S, New England BioLabs Inc.), and immediately thereafter the ligase was heat-inactivated at 65°C for 10 minutes.

The final concentration of DNA was then measured using a nanodrop, and found to be 140 ng/$\mu$L. As it was impossible to remove the unwanted DNA, not all of this is product however, and each lambda DNA was in a ratio of 1:20 with the handles. In this final product, one full lambda DNA molecule plus associated handles is the mass corresponding to one DNA tether as designed. The lambda DNA was in a ratio of 1:10 with the 500 bp handles, and so the mass associated with one construct is the mass of one lambda DNA molecule plus an additional 5 kbp. This notional mass is approximated by the formula [113]

$$\text{Molecular weight[g/Mol]} = 607.4N + 157.9$$

where N is the number of nucleotides in the system. Lambda DNA is 48,502 base pairs in length [114] thus accounting for 97,004 nucleotides, while the 5 kbp associated with the handles gives another 10,000 nucleotides for a total of 107,004. Substituting this in to the above equation gives an approximate molecular weight of $6.5 \times 10^7$ g/Mol. 140 ng/$\mu$L is equivalent to 140 mg/L and therefore the concentration of the designed DNA constructs is

$$\frac{140 \times 10^{-3}}{6.5 \times 10^{7}} = 2.15\,\mathrm{nM}$$



(a) 0.5% agarose gel showing the attachment of the 500 bp handles

(b) 1% agarose gel showing the 15 kbp construct's interactions with antibodies

Figure 2.5: Agarose gels demonstrating the construction and activity of the 15 kbp tether. a) Wells from left to right: ladder line in the first column indicating 23.1 kbp; from top to bottom the 20 kbp, 14.6 kbp, and 13 kbp lambda DNA fragments from the digestion; the same fragments after ligation to the 500 bp handles. All three bands shift indicating that handle ligation has happened to each fragment, and the lack of any visible brightness where the unligated bands should be indicates that this process is highly efficient. b) Wells from left to right: 3 $\mu$L 15 kbp construct; 15 kbp construct + 2 $\mu$L 1 mg/mL Pierce Avidin (21121, ThermoFisher Scientific); 3 $\mu$L 15 kbp construct + 2 $\mu$L 100 $\mu$g/mL antidigoxigenin. The shifts in brightness indicates correct antibody activity

Once again, the DNA construct was tested with gel electrophoresis. Initially, a 0.5% agarose gel was run for 22.5 hours at 20 V to demonstrate the handles attaching as expected to the midsection. As can be seen in Figure 2.5a, the bands are all shifted upwards, corresponding to the handles attaching to the midsection and the 13 kbp and 20 kbp DNA offcuts. The absence of unshifted bands indicates that the process is highly efficient, though the smearing of bands precludes intensity measurements to quantify this. The full product was then run on a 1% agarose gel at 100V for 1 hour against a commercial ladder to demonstrate affinity for the antibodies. Once again SYBRSafe was used to image the DNA.

Figure 2.5b shows the bands in the expected places for the construct, with the DNA in the third occupied lane remaining in the well due to interactions with avidin. The construct bands in the fourth lane are dimmer than the control lane, and some brightness can be seen shifted to the well due to its interactions with antiDIG. The two gels taken together indicate that the procedure followed to create this construct was a success.

## 2.3 Imaging

### 2.3.1 Optical design

Brightfield experiments on nucleic acids have been a fruitful endeavour for many years, for example characterising DNA's response to added twist [92] and observing RNA polymerase stepping along base pairs [115]. However, to push understanding further it is necessary to beat the diffraction limit and image individual molecules to subpixel resolution. To this end, a bespoke dual-colour excitation and complementary imaging path were created by Dr Helen Miller and coupled to the optical tweezers.

#### 2.3.1.1 Design of the fluorescence excitation path

In general, the super-resolution work performed with DNA in the group is done in a method analagous to stochastic optical reconstruction microscopy (STORM) [116]. Fluorophores which are either photoswitchable, or for which exciting and emitting is a stochastic process, are used, both of which have two states: a long-lifetime dark state in which the fluorophore does not emit light, and a short-lifetime bright state in which it is emitting. Across a population of fluorophores under excitation, the majority will therefore be in their dark state. In STORM, with stochastically blinking dyes as used in this work, a high laser power is used to as quickly as possible push the fluorophores into their dark states. From this dark state, they then enter the bright state before turning dark again – an event known as blinking. In order to find the position of the fluorophores to subpixel resolution, the blinking events are tracked in the images taken. An approximate form of the point spread function of the fluorophore, usually a 2D Gaussian profile, is fitted to the individual blinking events, and the centre of the fit is taken to be the location of the dye molecule. The long-lifetime dark state is especially useful here. In order for the separate point spread functions to be individually resolvable, they need to be separated by at least the Rayleigh criterion:

$$d = \frac{1.22\lambda}{\mathrm{NA}} \tag{2.1}$$

where $\lambda$ is the emission wavelength and $NA$ is the numerical aperture of the objective lens. The dyes being mostly dark means the blinking events are likely to be sufficiently separated that the individual point spread functions can be resolved.

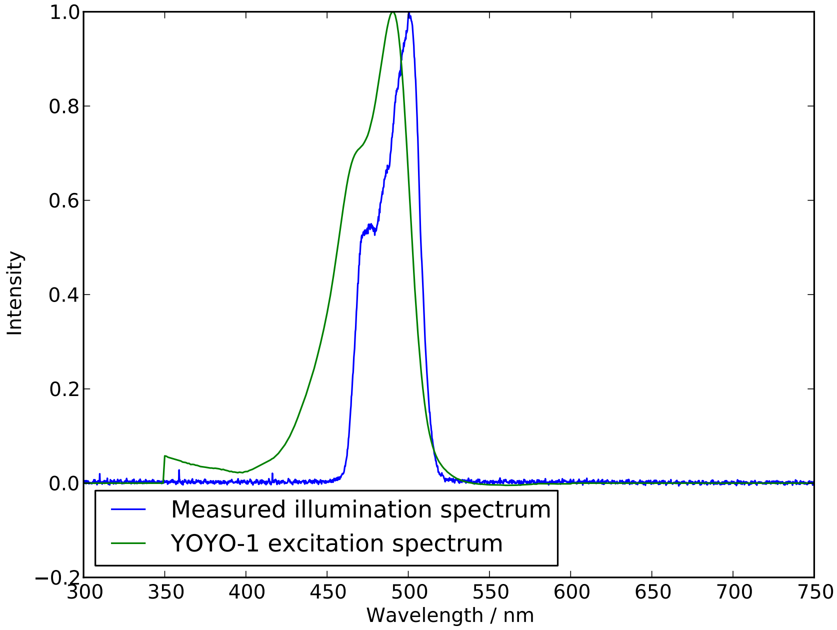


Figure 2.6: Excitation profile of YOYO-1 and the spectral properties of the corresponding excitation beam, normalised to the maximum value in each case. Data for YOYO-1 taken from [117]

The laser used is a super-continuum white light Fianium SC400-4, capable of emitting 4 W in total. For a dual colour microscope, two excitation wavelengths are selected by splitting the beam in two and passing each half through a linear gradient filter set. This results in two beams, one centred around 615 nm for excitation of the Tex615 fluorophore, and one centred close to 491 nm, the excitation peak of the DNA intercalating dye YOYO-1 which was used for this work. The **blue** excitation laser spectrum and its overlap with the YOYO-1 excitation spectrum is shown in Figure 2.6, which shows that the excitation profile is suitable for use with YOYO-1. Having passed through the linear gradient filters, the beams are combined again and pass through the remaining optics as one beam.

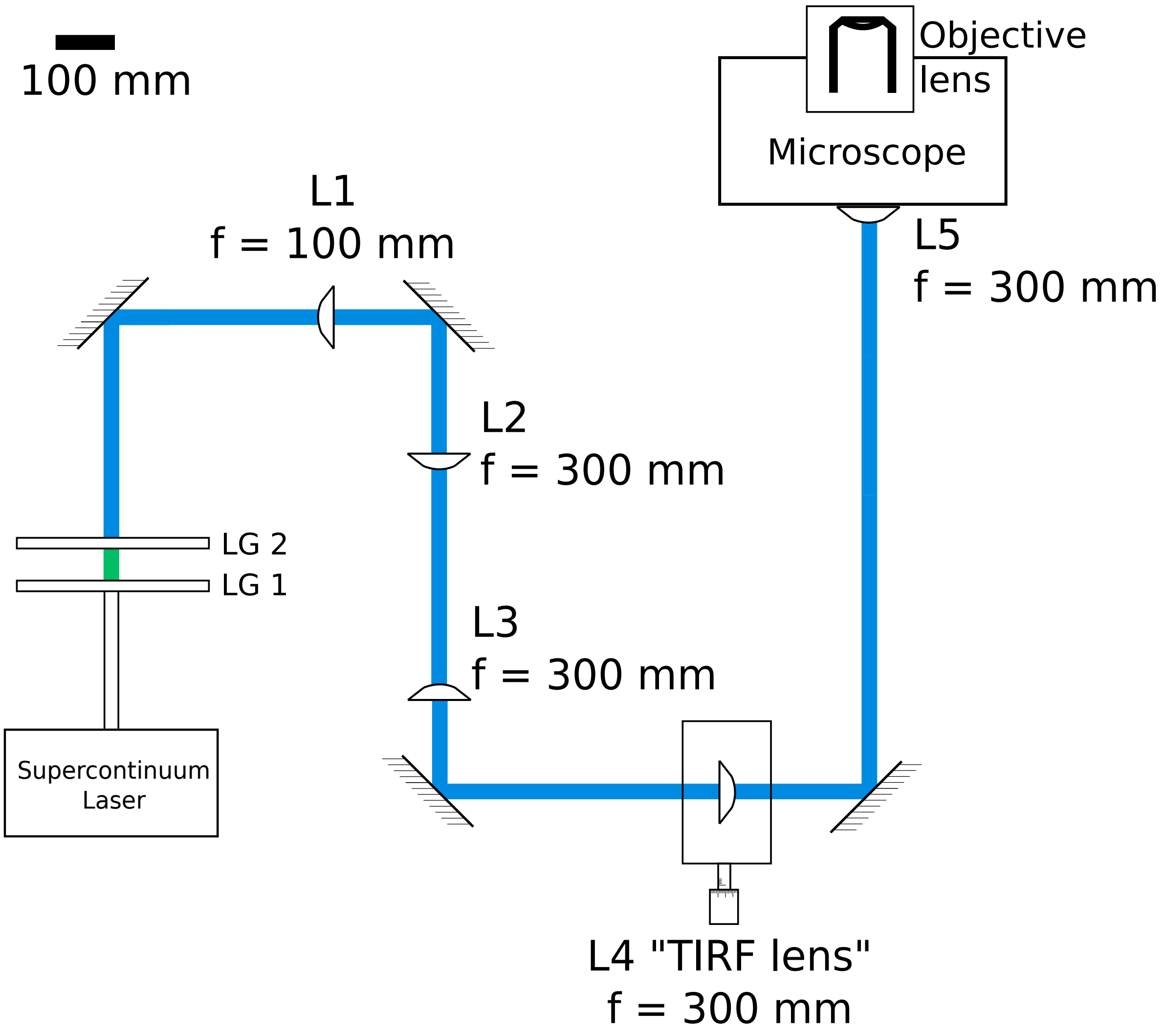


Figure 2.7: Schematic drawing of the blue fluorescence path. The correct wavelengths are selected from the white light from the super-continuum laser using two linear gradient filters labelled LG1 and LG2. The beam is then expanded 3× using a telescope consisting of one 100 mm and one 300 mm lens (L1 and L2 respectively) sited 400 mm apart. Following that is a second telescope consisting of two 300 mm lenses (L3 and L4) 600 mm apart. This achieves a 1:1 expansion, but the second of these lenses (indicated by “TIRF lens”) is mounted on a micrometer mount and can be translated laterally, which has the effect of selecting epifluorescence, HiLo, or TIRF illumination. A final lens (L5) forms a telescope with the objective lens and creates a beam with a measured full width at half maximum of approximately 25 $\mu$m at the sample when in epifluorescence mode.

As mentioned previously, it is desirable to put as much laser power into the sample to force dye molecules into their dark state. To facilitate this, the beams are expanded as minimally as possible, with a single expansion step of 3× in the excitation path. The beams then go through a 300 mm lens mounted at the back aperture of the microscope, which forms a telescope with the objective lens so that the beams exit the objective collimated and vertical for epifluorescence microscopy. The lens/objective lens telescope reduces the beam diameter significantly, and the beam size at the sample is found to be 25.8 $\mu$m FWHM. The beam is however not perfectly collimated, though this has little effect over the length scale of the sample depth – the microscope chamber is around 100 $\mu$m in depth. The maximum power which can be applied to a sample is 9mW giving a power density of 1800 mW/cm$^2$. This power is extremely high, and the beam is commonly attenuated in the excitation path with neutral density (ND) filters.

A schematic diagram of the excitation path, including the path of the emitted light, can be seen in Figure 2.7.

#### 2.3.1.2 After excitation: the emission path

Having excited the fluorophores in the sample, it is now left to collect the emitted light, magnify it further, and capture an image. The optics that accomplish this are collectively known as the emission path.

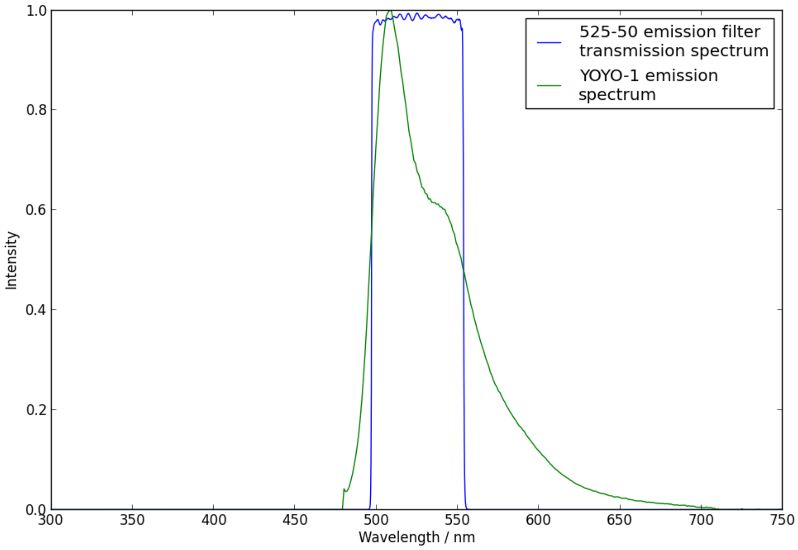


Figure 2.8: The emission spectrum of YOYO-1 and the transmission properties of the relevant region of the emission filter employed. YOYO-1 data from [117], emission filter data from [118]

The objective lens with its high numerical aperture is able to efficiently collect the emitted light, but it captures also any light from the excitation laser which has reflected from the coverslip. This rogue excitation light must be filtered out and so a

Semrock Brightline FF444/521/608-Di01 emission filter was inserted underneath the dichroic mirror. The emission filter, the spectral properties of which are seen in Figure 2.8, excludes all wavelengths other than those from the target fluorophores. As the microscope is constructed explicitly to be dual-colour, the emission filter has multiple transmission peaks such that both red and green fluorophores may be imaged, however only the salient region for YOYO-1 is shown in Figure 2.8.

The light now cleaned appropriately, it leaves the microscope aperture. The microscope as supplied contains a tube lens within the microscope body which makes this aperture a conjugate image plane to the sample, and as the objective lens magnifies the image by 100×, a camera could be placed here and the sample imaged immediately. However for higher resolution and super-resolution it is necessary to increase the magnification further. To this end, the emitted light is then expanded using another telescope composed of a 75 mm and 150 mm achromatic lens pair which magnify the image by 2X.

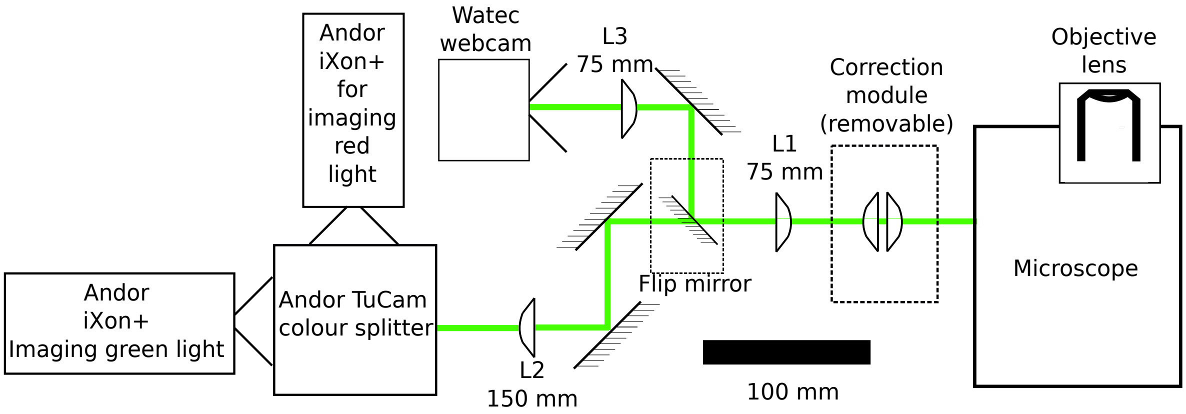


Figure 2.9: The microscope emission path. During a trapping experiment, the light initially passes through the focus correcting module. Following that the light passes through a 75 mm lens. A flip mirror (marked inside a dashed box) selects the Andor iXon+ cameras or the Watec webcam which has a larger field of view. If the Andor cameras are selected, the light passes into a 150 mm lens, expanding the beam by 2X, and subsequently an Andor Tucam module which splits the light according to wavelength and focusses the red light on to one camera and the green light on to another, and in this way two colour imaging is possible. If the Watec webcam is selected, the light meets another 75 mm lens and the beam is therefore not expanded. The optics within the microscope body itself including the dichroic mirror and tube lens are not shown.

The light is then ready to be imaged. It is directed into an Andor TuCam dual-camera module which splits the light with a dichroic mirror such that the red light is focussed on one camera and the green focussed on the other. A filter module within the

TuCam also removes any remaining infra-red light. The TuCam was selected as it simplifies the placements of cameras tremendously – alignment of the cameras themselves with the dichroic mirror is largely taken care of.

The cameras used were both Andor iXon+ EMCCD with chips of 512×512 pixels, a maximum absolute EM gain of 300, and air-cooling capability to -80℃. The microscope body referred to in this section is an inverted commercial Nikon Eclipse Ti-S, which has the advantage of standardised sizing for additional optical components, a dichroic filter wheel, an objective lens selector, and compatability with aftermarket mechanical equipment such as the piezoelectric nanostage which will be discussed in due course.

The total effect of this imaging path is to allow the microscope to capture images at 80 nm/pix, with superresolution data localised to around 30 nm reported using this microscope [119]. The schematic of this emission path is found in 2.9.

### 2.3.2 Choice of fluorescent dyes

Many fluorescent dyes have been produced in the decades since fluorescence imaging made its breakthrough. However, few are well suited to the application in this thesis. Specifically, proteins such as the green fluorescent protein (GFP) can be cumbersome to work with as they require expression and attachment to the biological sample to be imaged, while their molecular weight may also significantly disrupt the mechanical processes of the contour of DNA. Dyes which must be attached to the sample are largely useful in the context of DNA for tagging and identifying individual, small regions of interest, for example labelling the ends of a construct. For imaging the full contour, a dye which binds uniformly to the DNA, and which can be added to the sample at the time of the experiment is preferable.

Rather than tagging, therefore, small chemical dyes are used for imaging nucleic acid constructs. These interact with the DNA in one of three primary ways – major groove binding, minor groove binding, and intercalating. A schematic of these three binding modes is presented in Figure 2.10.

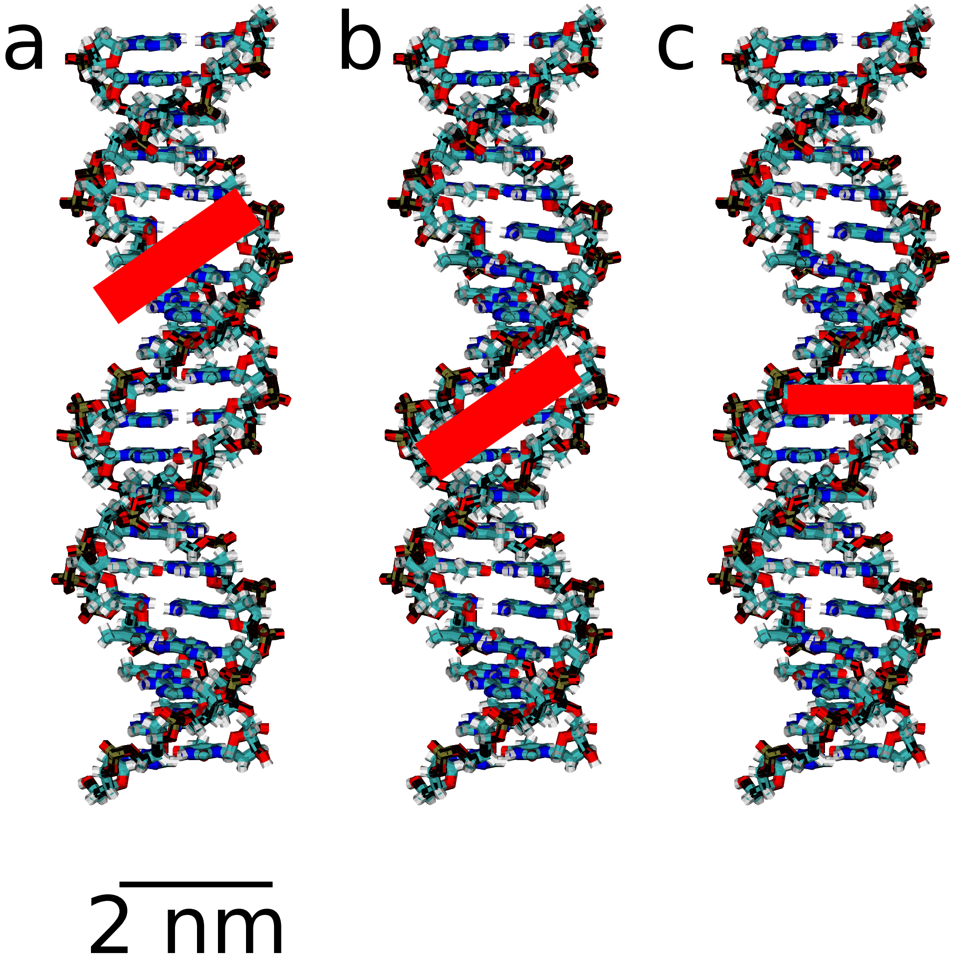


Figure 2.10: Three possible dye binding modes. a) Major groove binding; b) minor groove binding; c) intercalating. The bound molecule in red is not drawn to scale.

For the ultimate aim of this work – super-resolution imaging of the full contour length of DNA – a candidate dye must meet certain criteria. First, it should have a low fluorescence background, as the signal is likely to be weak and the signal-to-noise ratio (SNR) must be optimised from the start. Secondly, it should have properties commensurate with super-resolution imaging – that is to say, it should blink when under excitation. Thirdly, it should damage the DNA as little as possible. It should also bind to the DNA in an indiscriminate way to ensure the full contour is illuminated.

Many cyanine dyes such as thiazole orange (TO) [120], oxalose yellow (YO), YOPRO, and YOYO-1 [120] have been successfully used to image full contours of DNA over significant timescales. One of the most common dyes used to gather super-resolution information from DNA is the bis-intercalator YOYO-1. Used to great effect in the early work of DNA and optical tweezing [104], the dye has also been shown by various groups to allow subpixel location of individual fluorophores, for example [119].

YOYO-1 also has fluorescence enhancement behaviour whereby the brightness of the fluorophore increases by three orders of magnitude upon binding to DNA [120], a property which should reduce the overall background and give good signal to noise ratio. However, it is known to be damaging, and mitigation of this is a great concern. YOYO-1 is also quite a disruptive dye. It is known to extend the contour length of DNA by around 38% and to cause underwinding, but not to change the persistence length [85]. Its high affinity and low off-rate mean that it may disrupt the activity of proteins or molecular machines associated with DNA, though that remains to be seen.

Other dyes which bind the DNA in less disruptive (i.e. minor groove-based) ways are also on the market, and one which is a good candidate for super-resolution work is SYTO-13, a green dye in the SYTO family. This is a minor groove binding dye which exhibits good photoblinking under sustained excitation (see [119]), and its activity is not sequence dependent. Its minor groove binding properties may also mean that when investigating the interactions between DNA and motor proteins, the dye may pose less of a roadblock to the molecular machine. However, it is significantly less bright than YOYO-1 and does not experience a fluorescence boost upon binding, so free dye in the imaging buffer will lead to a poor SNR.

Due to its fluorescence properties and brightness, YOYO-1 was chosen. The work in this thesis does not deal with motor proteins or indeed any molecule other than DNA itself, and as a result the dye's disruptive properties are less problematic than they may be in other assays. Its brightness, robustness, and relative cheapness mean that it is a good dye to begin with as optimising imaging conditions for it should be less challenging than doing so for a dimmer, more condition-dependent dye. However, its capacity to produce free radicals which attack the DNA backbone of significant concern and as a result steps must be taken to prevent DNA damage.

#### 2.3.2.1 Keeping the lights on

Under excitation, most dyes photobleach, and in so doing emit free radicals, usually reactive oxygen species (ROS). These ROS are extremely damaging to biological samples, and can easily create breaks in the covalent bonds which make up the backbone (nicks). Multiple nicks close together on the molecule will cause a catastrophic snapping event, an example of which is seen in Chapter 3. These ROS are seen *in vivo* also and the cell has multiple chemical mechanisms to prevent DNA damage. For *in vitro* experiments, ROS scavenging systems are used to reduce damage to the sample.

In the early days of YOYO-1 usage, a free radical scavenging system consisting of catalase, glucose oxidase, and glucose was used [104]. The removal of oxygen increased the blinking rates of the dyes unacceptably, however [121], and so this was suppressed through the addition of triplet state quenchers such as β-mercaptoethanol or Trolox, which works to quench the triplet state through electron transfer followed by a complementary redox reaction [122].

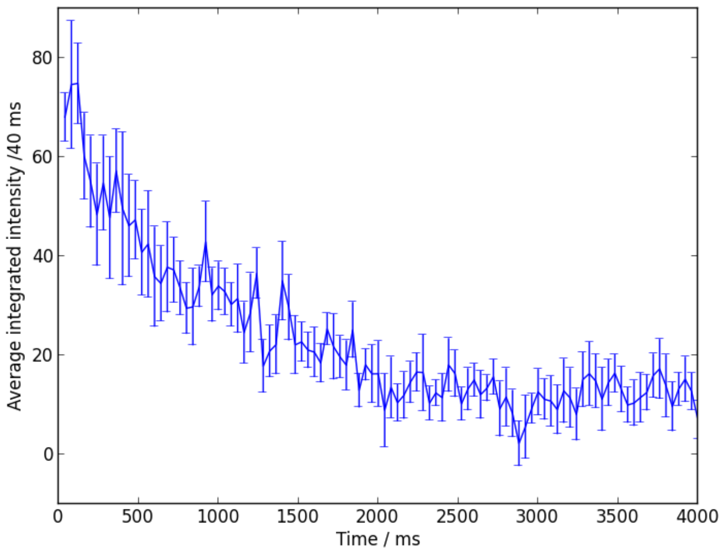
As in [123], the free radical scavenging system selected in the first instance for fluorescence imaging was 2 mM Trolox, 6% glucose, 1 mg/mL glucose oxidase and 40 $\mu$g/mL catalase.

However, preventing DNA damage is the first step only. It is inconceivable that all ROS will be immediately removed from the system before having a chance to cause damage; further, mechanical shearing from pipette tips can cause nicks in the DNA. Any damage that *is* caused needs to be repaired, as nicked DNA is no longer torsionally constrained, but is free to rotate around the remaining backbone's phosphodiester bond at the site of the nick. In the cell, DNA ligase works to repair such nicks to form unbroken strands of DNA, and it is used for the same purpose in experiments. The imaging buffer was therefore supplemented with T4 DNA ligase (New England Biolabs) in the concentration 4 units/uL (where a unit is defined as the amount that gives 50% ligation of HindIII fragments in 30 minutes at 16℃ in a total volume of 20 $\mu$L in 1X T4 ligation buffer [124]) to seal any nicks that may form. The provided NEB T4 Ligase Buffer recipe was used as it has been optimised for the ligase enzyme itself. It consists of 50 mM Tris-HCl at pH 7.6, 10 mM $MgCl_2$, 10 mM DTT, and 1 mM ATP at 1× concentration [125].

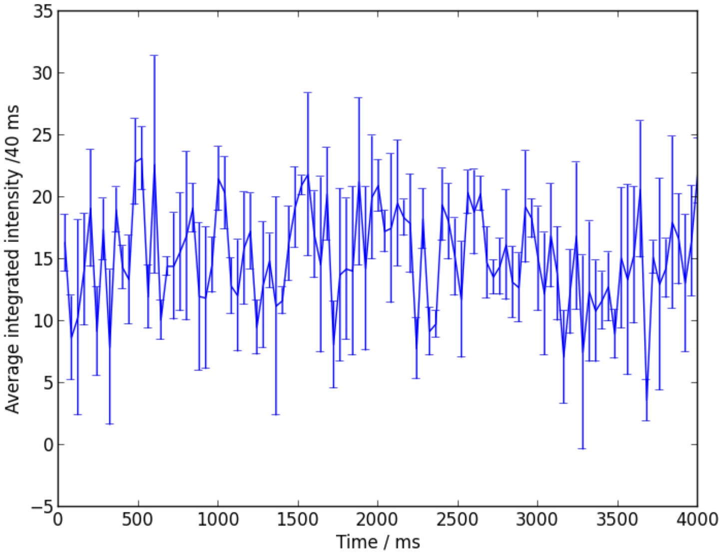
#### 2.3.2.2 The Fluorescence Imaging Buffer

As noted in Section 2.2.2.1, previous studies have used an oxygen scavenging system consisting of 6% D-glucose, 1 mg/mL glucose oxidase, 2 mg/mL Trolox and 40 $\mu$g/mL catalase [123]. However, how well this works in the context of column-purified DNA/YOYO-1 conjugates is uncertain.

To investigate, two sandwich assays were performed, in which once again DNA was dropped on to a microscope slide surface and covered with a plasma-cleaned cover slip. One sample was DNA with YOYO-1 in the concentrations 8.6 pM and 500 nM respectively in a buffer of 40 mM Tris-HCl + 10 mM $MgCl_2$, while the other had the same concentrations of DNA and YOYO-1 in a buffer with the full free radical scavenging system, composed of 40 mM Tris-HCl, 10 mM $MgCl_2$, 10 mM DTT, 1 mM ATP, 10 units/mL T4 DNA ligase, 2 mg/mL Trolox, 1 mg/mL glucose oxidase, 6% D-glucose, and 40 $\mu$g/mL catalase.



(a)



(b)

Figure 2.11: Average intensity per 40 ms exposure through time for a) DNA+YOYO-1 without the oxygen scavenging system and b) with the oxygen scavenging system present. The number of molecules per data point was 5, and the error bars represent the standard error in the mean of these 5 molecules.

The DNA was imaged with laser power measured to be 1.3 mW at a wavelength of 491 nm. Acquisitions were 4 s long with an exposure time of 40 ms. The constructs were seen to be approximately circular compact structures instead of stretched out on the cover slip surface as in [119], presumably due to their shorter length. A circular area containing one putative DNA molecule was defined and its total fluorescence intensity through time found with ImageJ [126]. The same was done for an equally sized area adjacent to the molecule to estimate background intensity. The background was then subtracted from the DNA signal and the intensity plotted. This process was done for multiple (N>5) DNA molecules.

As can be seen in Figure 2.11, the brightness was significantly reduced for the sample imaged with the oxygen scavenging system present, contrary to expectations. It was hypothesised that the Trolox, a triplet-state quencher originally designed for use with Cy3, may be interfering with the proper fluorescence of YOYO-1. The glucose, glucose oxidase, and catalase together act as the free radical harvesting system, while the Trolox is to prevent photoblinking. Therefore, sacrificing the Trolox to increase fluorescence yield would be worthwhile if that were the problem.

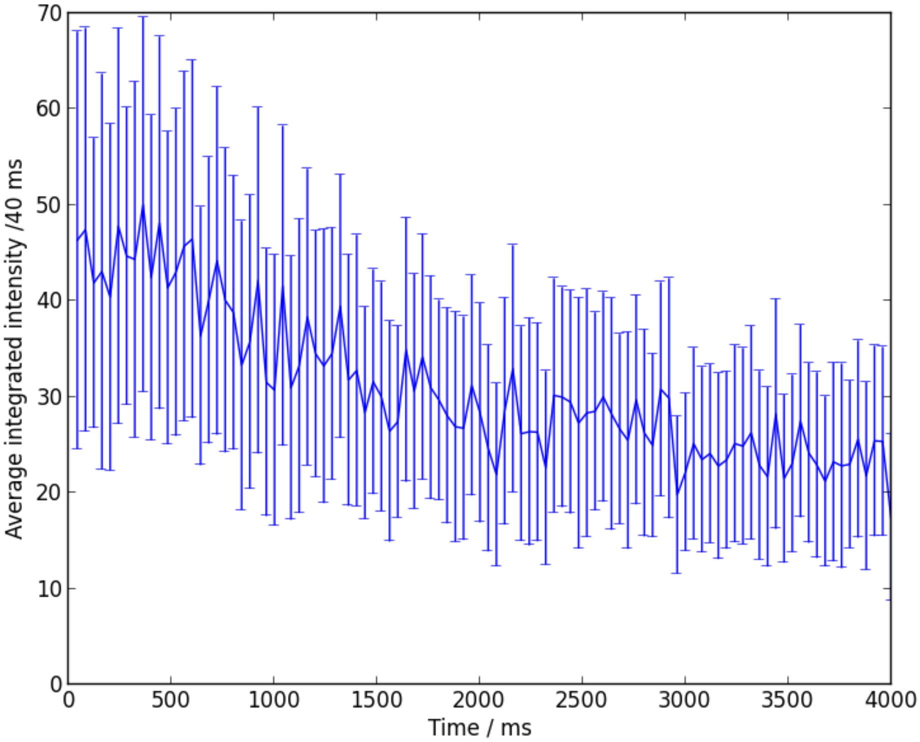


Figure 2.12: Average intensity per 40 ms exposure through time for DNA+YOYO-1 in an oxygen scavenger without Trolox. 5 DNA molecules were images for each data point. The point is the average of the 5 molecules' intensities and the bars are the standard error in this mean with N=5.

A third sandwich assay was carried out, with the buffer containing the free radical scavenger without Trolox present. This sample was analysed in the same way as in Figure 2.11, and the results are given in Figure 2.12. The brightness here has increased, though not to the same levels as in Figure 2.11a, suggesting that the Trolox was participating in the observed loss of fluorescence was not the only driving force. However, as the intensity was raised on its removal to easily visible levels, Trolox was not used for attempted fluorescence imaging assays.

### 2.3.3 Data acquisition software and settings

A great deal of data is produced by even one experiment using such multifaceted equipment and techniques, and saving, storing, and processing this requires specialised software.

The images, both brightfield and fluorescence, were taken using the Andor Solis package [127], a commercial imaging and analysis program which is provided by Andor for use with their cameras. All images were taken with an exposure time of 40 ms, frame transfer was enabled so that frames were continuously acquired, and for fluorescence images the electron multiplying gain was enabled and set to its maximum value. For all images acquired, the emCCD chips were air cooled to -80℃.

The cameras themselves were connected to shutters in the excitation path which opened when an acquisition began so that the sample was illuminated only when images were recorded. The camera fire output which indicates when a frame was being captured was also recorded and used to synchronise data, as will be explained in more depth in due course.

The other data sources which need recording in this setup are as follows: the QPD $V_x, V_y$ and $V_{sum}$ readings, the measured voltage across each pair of Helmholtz coils, the position of the piezoelectric nanostage, and the camera fire signal as mentioned above. Each of these is readily measurable as the hardware has analogue measurement ports to connect to. These ports are then wired into a commercial 6 channel DAQ card (National Instruments, CompactDAQ C Series) which can acquire data at a rate of 10kHz and stream it to the computer's hard drive via a simple LabVIEW interface. In this way data can be captured every 0.1 ms, and as the camera fire output is recorded, the data from magnetic tweezers or the nanostage can be correlated with captured images.

This method has many distinct advantages over reading the data to the computer and synchronising through timestamps and software. Most importantly, Microsoft Windows offers little guarantee that a timestamp is correct. Two pieces of data taken at the same time may experience different lag while being saved and the apparent timestamp may therefore differ by tens of milliseconds. The other consideration to be made is that each function of the computer requires resources such as RAM or CPU cycles, and requiring the computer to synchronise timestamps as well as control the nanostage and operate the force clamp may overwhelm the system. Taking data with as minimal manipulation and interference as possible and instead post-processing decreases the risk that the computer may exhaust its resources at a critical time and ruin an experiment. In this project, collection of data via DAQ card and LabVIEW utilising hardware synchronisation was found to be the only feasible way to get the precision and time correlation required.

The LabVIEW code which performs this acquisition has multitudinous other functions. It controls also the piezoelectric nanostage for fine manipulation of samples (a Mad City Labs Nano-LP300 with 300 $\mu$m range in all three axes and 0.6 nm resolution [128]), sends waveforms to the signal generators and hence controls the magnetic tweezers, and when in force clamp mode minimally processes the QPD data to find forces applied in real-time and moves the nanostage to keep applied force at a constant desired level. This process will be explained in more detail later in this Chapter.

## 2.4 Tweezers

Though there are many ways to optically and magnetically tweeze particles, only those relevant to the work presented in this thesis will be presented here. Specifically, the techniques used were the optical trapping of a large particle by a focussed laser beam and application of torque using Helmholtz coils driven by an alternating current.

### 2.4.1 Principles of optical tweezing

#### 2.4.1.1 The large-bead limit

Beams of light have two primary ways to exert a force on a particle they encounter. For a small dielectric particle, the gradient of the electromagnetic field exerts a force towards the peak of intensity of the beam – i.e. for a Gaussian profile the centre. For larger particles, it is the linear momentum of the light which is most relevant. Several review articles exist which examine optical trapping and its uses in depth, including [42, 129, 130, 131].

Two regimes of optical trapping exist in which the different forces dominate: the Rayleigh limit, in which the trapped particle is much smaller than the wavelength of trapping light and is trapped primarily through electromagnetic gradients, and the Mie regime which is the case for beads or particles which are larger than the wavelength of the trapping light [132]. In this large bead limit, when the diameter of the trapped particle is greater than the wavelength of the trapping beam, the contributions to the force from electric and magnetic fields may be ignored [132]. Instead, a purely ray-optics based explanation for the behaviour of the trap is most appropriate.

First, in order to generate an optical trap in this fashion, a collimated beam of monochromatic light is focussed to a diffraction-limited spot using an objective lens. A lens's numerical aperture is a measure of the range of angles at which light may be accepted or emitted by the lens, and is given by

$$\mathrm{NA} = nsin(\theta) \tag{2.2}$$

where $n$ is the refractive index of the medium the lens operates in and $\theta$ is the maximum half angle of the cone of light that can pass through the lens. For optical trapping, it is important to focus an expanded beam to as small a focus as possible, and so a high numerical aperture lens is chosen – a typical value being 1.49. The full width at half-maximum (FWHM) of the diffraction-limited spot, is approximated by

$$2w_0 = \frac{2.44F\lambda}{D} \tag{2.3}$$

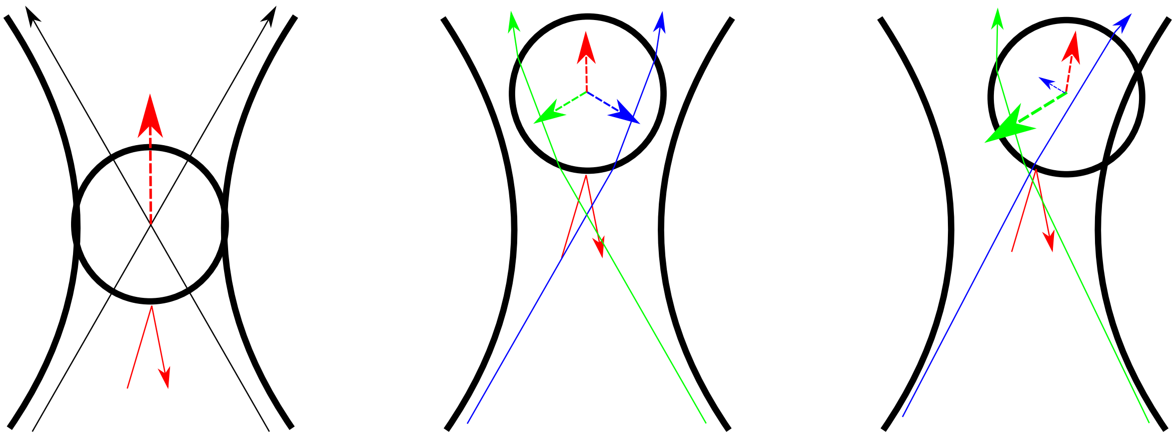
where $w_0$ is the radius of the beam at the beam waist, $\lambda$ is the wavelength, $F$ is the focal length of the lens, and $D$ is the diameter of the illuminated part of the focussing lens. Using the parameters for the microscope used in this work:

$$\begin{aligned} \lambda &= 1064\,\mathrm{nm} \\ F &= 2\,\mathrm{mm} \\ D &\approx 5\,\mathrm{mm} \end{aligned}$$

we obtain

$$2w_0 = \frac{2.44 \times 1064 \times 10^{-9} \times 2 \times 10^{-3}}{5 \times 10^{-3}} \approx 1038\,\mathrm{nm} \tag{2.4}$$

Optical traps are at their stiffest when the beam waist coincides with the size of the bead to be trapped – in this case a 1 $\mu$m bead would be most ideal if there were no other considerations to be made.


(a) A bead at the focal point of an optical trap. The bead experiences an overall upward force due to radiation pressue, the horizontal components cancelling by symmetry.

(b) A bead trapped in the minimum energy state of an optical trap. The bead experiences no force here

(c) A bead to the right of the central point of an optical trap. The horizontal components of the momentum transfer no longer are equal and as a result the bead experiences a leftward force.

Figure 2.13: Three possible beads positions in an optical trap and the resultant forces. Light rays are indicated by thin solid arrows and the resultant forces by thick dashed arrows. Light rays and the forces they produce are drawn in the same colour. Figure not drawn to scale.

When a bead is introduced to this focussed point, it experiences various forces. One is radiation pressure which pushes it in the direction of beam propagation due to the light reflecting from the bottom surface. The other forces are due to the change in momentum of the light as it exits the trapped bead, and the interplay between these forces governs the behaviour of the system.

Figure 2.13 shows the forces on a bead due to a focussed laser beam, and it can be seen that as the bead moves to the right there is a force to the left, and *vice versa*. Any local smooth, continuous, and continuously differentiable function may be approximated by a Taylor expansion, and for a Gaussian this is approximately quadratic for small displacements from the maximum. The centre of the trap is such a local minimum and as such in the centre both the intensity profile and the potential energy experienced by the bead may be approximated as parabola. As the force experienced in the centre of the trap is proportional to the gradient of the laser intensity, the force applied is thus approximately linear in this central region and can be described well using Hooke's law $\vec{F} = -k\vec{s}$ where $\vec{F}$ is the force applied, $k$ is the stiffness of the trap, and $\vec{s}$ is the vector displacement from the centre of the trap.

#### 2.4.1.2 Practical considerations

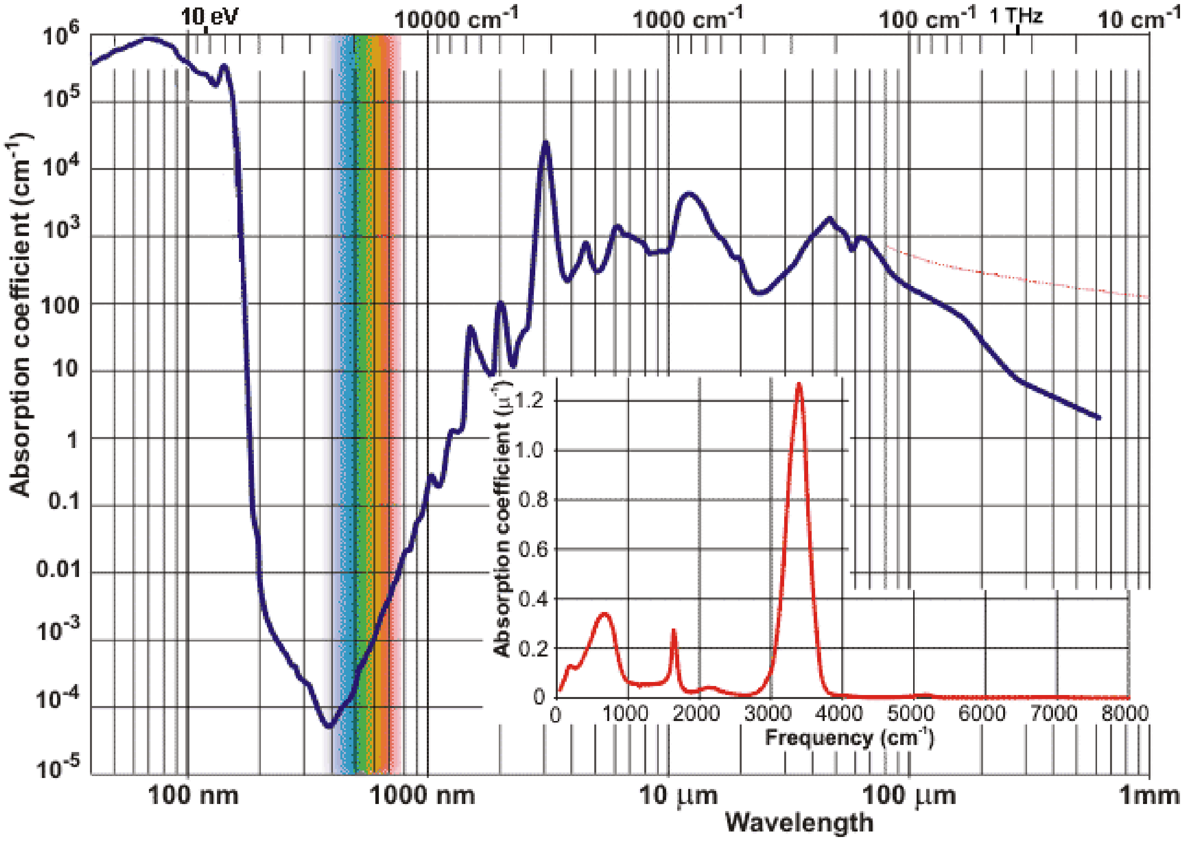


Figure 2.14: Absorption spectrum of water. Image taken from http://www1.lsbu.ac.uk/water/water_vibrational_spectrum.html Retrieved 27th August 2018.

In order for the above treatment to be valid, certain conditions must be met. The first is that the beam profile should indeed be Gaussian – in practice, only the centre of a laser beam will truly fit this profile. Therefore when designing laser tweezers the beam is traditionally expanded so that it overfills the back aperture of the microscope [129] so that only the central part of the beam, which has the best spatial profile, is then focussed by the objective lens. However, recent work has shown that underfilling the back aperture can result in higher trap stiffness [133], however in this work the more standard overfilling technique was used. In either case, it is required that the light is collimated when going into the objective lens, which must be ensured during building.

The choice of trapping wavelength is also an important consideration. Water absorbs a great deal of light, as can be seen from its absorption spectrum in Figure 2.14. However, there is an order of magnitude reduction in the absorption of light by water in the region 750-1200 nm. For this reason a great many optical traps are built using

1064 nm lasers, as this also happens to be a convenient wavelength for manufacturers. A secondary benefit of using a 1064nm laser in this work is that the dyes used have excitation profiles far away from this value, so that except by a small amount of two-photon excitation the trapping laser should not photobleach the sample before it can be imaged properly.

### 2.4.2 Design and construction of the optical tweezers

As discussed above, the comparatively low absorption by water of wavelengths in the range 750-1200 nm makes 1064 nm a natural choice of wavelength, and in common with most other laser tweezer setups the optical tweezers presented here are built around a 1064 nm laser, specifically a Laser Quantum Ventus capable of generating 4 W of 1064 nm laser light. This maximum power is far too high to be used for a single trap – the water would boil, the trapped bead would explode, and any biological samples would be ruined instantly. To attenuate the power, the laser first passes through a half-wave plate, the angle of which changes the fraction of light which is horizontally and vertically polarized. A polarisation beam splitter separates the two components, one of which is safely removed to a beam dump. In this way the power can be easily varied using a PC-controlled mechanical mount on the half-wave plate, and the power can be reduced to safe levels for construction of the tweezers. Mirrors and lenses are used which have an infrared coating applied to minimise loss and aberration.

Having been suitably attenuated, the beam is now ready for expansion. The expansion telescope is composed of a 100 mm lens and a 300 mm lens spaced 400 mm apart for a total expansion of 3×. The threefold expansion of the trapping beam makes it on the order of 1 cm in diameter, which is enough to overfill the back aperture of the microscope. At the back aperture of the microscope is a 300 mm lens used to form a telescope for the fluorescence excitation laser, and it is therefore necessary to place another 300 mm lens 600 mm away from the back aperture to form a 1:1 telescope and ensure the trapping laser is collimated as it enters the objective lens. As discussed previously, the objective lens takes this collimated beam and focusses it to a diffraction-limited spot after 2 mm, creating the trap itself.

The objective lens used for this work as a Nikon 100X CFI Plan Apochromat oil immersion lens with a numerical aperture of 1.45.

A diagram of the optical path of the tweezer is shown in Figure 2.15.

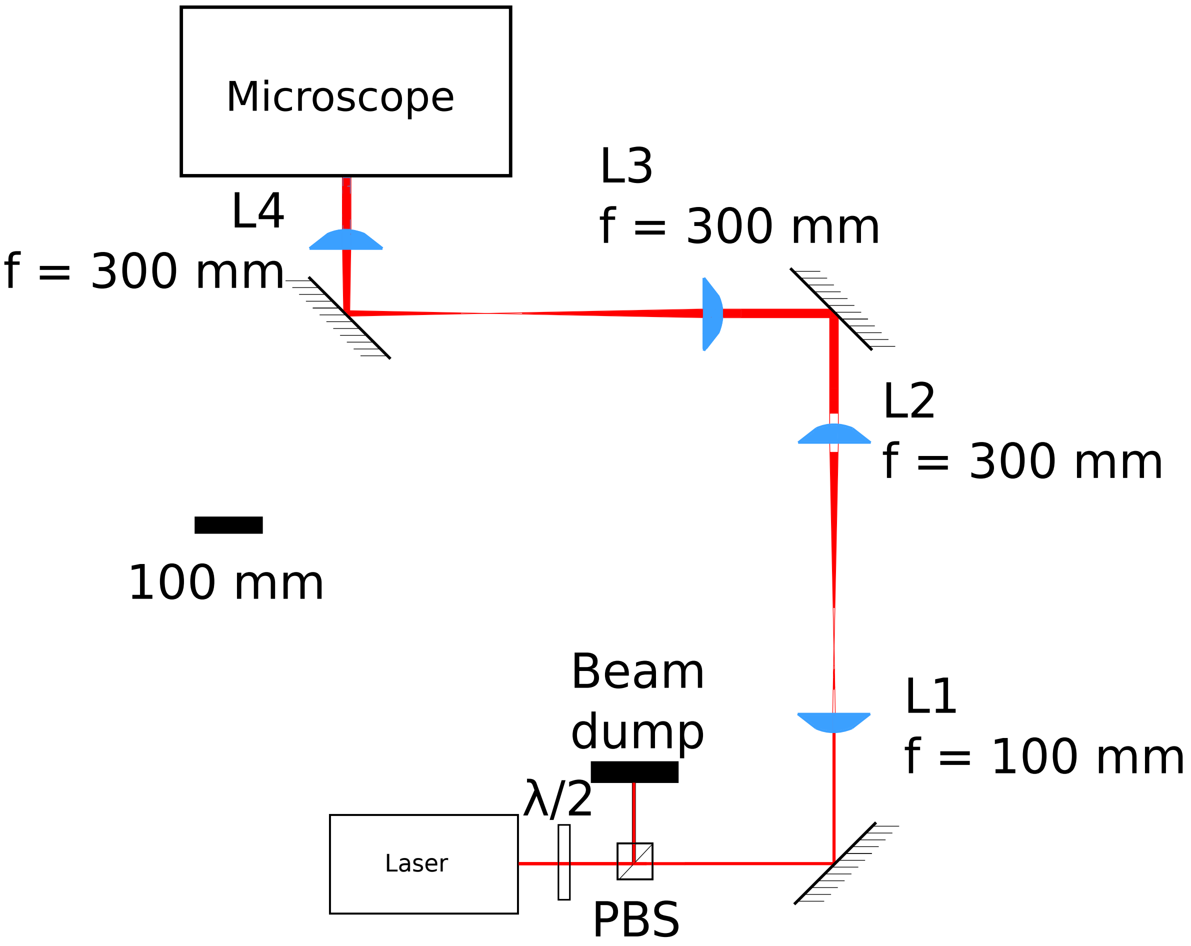


Figure 2.15: Schematic of the optical path for the laser tweezers. Lenses L2, L3, L4 are 300 mm focal length, and lens L1 is 100 mm focal length. The beam is attenuated by a rotating half-wave plate marked $\frac{\lambda}{2}$ and subsequent polarisation beam splitted marked PBS. Excess light is safely removed to a beam dump. The trap path then consists of two telescopes, one with a 3× expansion to overfill the objective, and one which is 1:1 to ensure the beam is collimated entering the back aperture of the microscope.

### 2.4.3 Measuring Forces with the QPD

Measuring the force applied by a spring requires two pieces of information: the displacement and the trap stiffness. Both of these may be interrogated using back focal plane interferometry [78], an extremely powerful tool which in this case makes use of the trapping laser itself, though the same method may be identically applied with a second, low-power, tracking laser.

After interacting with the bead, the trapping laser produces a characteristic interference pattern, which is collected by the condensing lens and focussed on to a quadrant photodiode (QPD). The quadrant photodiode (QPD) is a cheap, high-bandwidth (up to hundreds of kHz [134]), effective measurement tool. Four p-n semiconductors are arranged in separate quadrants. Incident light on the sensor excites electrons at the p-n junction, causing some to enter the conduction band and creating electron-hole pairs. Depending on the device, either a current or a voltage is produced and measured. As is usual for a semiconductor, the voltage produced is a function of the total intensity hitting each individual photodiode.

As the bead moves in the trap, the interference pattern changes and as a result the voltage produced by the photodiode changes in a predictable way depending on the location of the bead. If the response of the QPD is well characterised, the displacement can then be trivially calculated from the measured voltage. The QPD allows for this process to happen independently in both vertical and horizontal directions so that the 2D position of the bead relative to the trap centre can be found. Additionally, measurement of the total intensity of light reaching the QPD can be used to find the displacement along the $z$ axis [129].

The four photodiodes work together as four pairs: a top pair, a bottom pair, a left pair, and a right pair. The QPD output is thus split into three readouts: $V_x$ and $V_y$, which are the differential voltages produced by the left and right pairs and top and bottom pairs, respectively, and $V_{sum}$which is the sum of the voltages produced by each quadrant. If the bead is on the right hand side of the trap, the $V_x = V_L - V_R$ will increase, as the shadow of the bead is cast on the right hand side of the QPD, producing less voltage in that half of the instrument. The same argument applies for $V_y$, while $V_{sum}$ is used for normalising the QPD data, ensuring that different heights in $z$ of the bead and hence different intensities hitting the QPD, do not affect the measured positions in $x$ and $y$ [129]. Knowing the position of the bead in the trap to high accuracy both spatially (resolutions of 0.1 nm have been reported [115]) and temporally (up to hundreds of kHz [134]) means that the spring constant of the trap can be found via the robust and accurate power spectrum method. This method decomposes the bead's trajectory into characteristic frequencies, and using that information to deduce the strength of the notional spring in the centre of the trap by fitting it to a Lorentzian curve.

First, QPD data is taken of an untethered bead trapped in the optical tweezer. Each axis is treated as independent, so the following procedure is undertaken for the voltage readings in both the $x$ and $y$ directions. The collected data is initially Fourier transformed using the fast Fourier transform (FFT) algorithm, and the frequency data is plotted on a log-log scale, as seen in Figure 2.16.

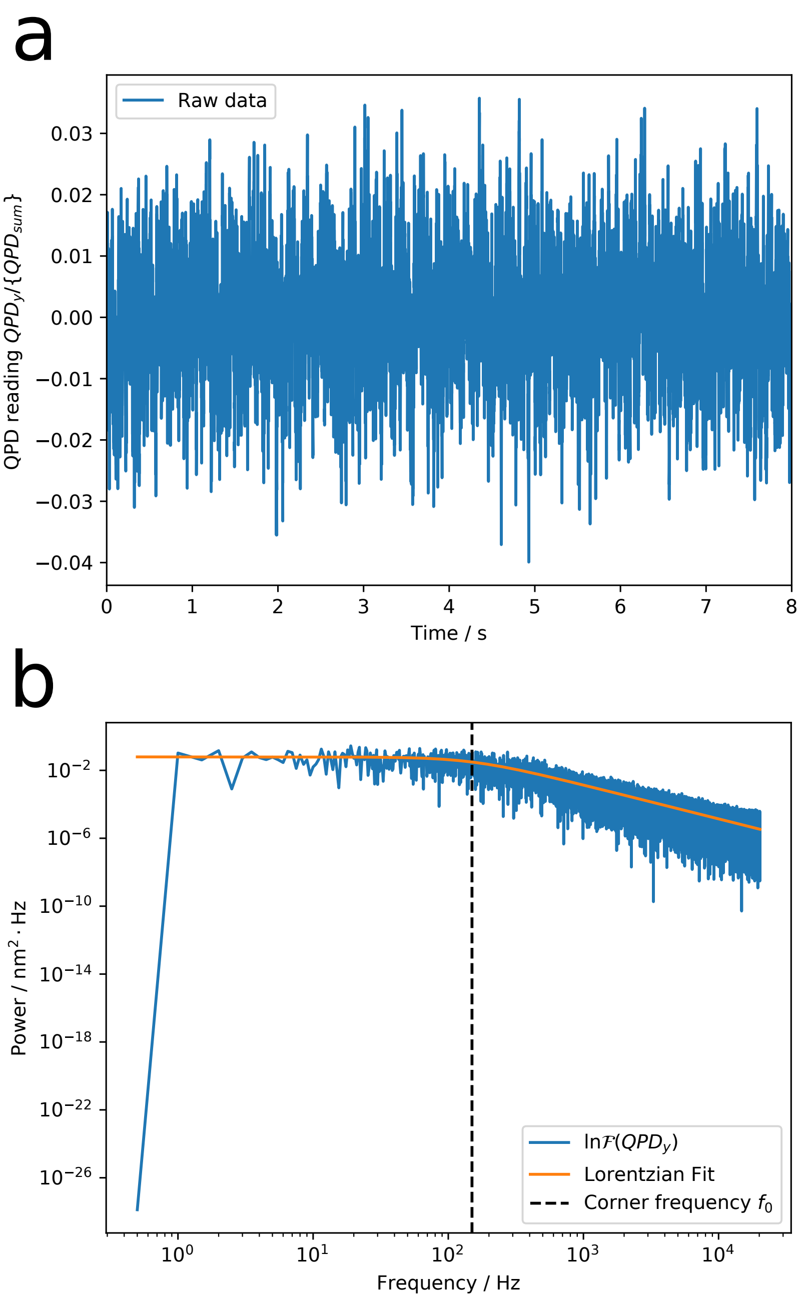


Figure 2.16: Stages of the data analysis to generate trap stiffness in $y$. a) Raw data; b) Fourier transformed data plotted on a log-log scale with Lorentzian fit. The corner frequency $f_0$ is indicated with a dashed line

This data is now ready to be fitted. The Langevin equation for a bead trapped in an optical trap is

$$F(t) = m\frac{\mathrm{d}^2x}{\mathrm{d}t^2} - \gamma\frac{\mathrm{d}x}{\mathrm{d}t} - kx \tag{2.5}$$

where $F(t)$ is the force on the bead at time $t$, $x$ is the displacement from the centre of the trap, the hydrodynamic drag $\gamma$ for a fluid of viscosity $\eta$ and bead radius $r$ is given by $\gamma = 6\pi\eta r$, and $k$ is the trap stiffness. The inertial term $\frac{d^2x}{dt^2}$ may be safely ignored due to the low Reynolds number of the system, leaving

$$F(t) = -\gamma\frac{dx}{dt} - kx \tag{2.6}$$

The derivation of the power spectrum expected for a system of this kind has been performed previously and is omitted here – the mathematics may be found in many publications including [111, 135]. The power spectrum for a harmonically trapped bead is given by

$$PSD(f) = \frac{2k_BT\gamma^2}{k^2 + (\gamma f)^2} \tag{2.7}$$

Using the relation

$$k = 2\pi\gamma f_0 \tag{2.8}$$

where the corner frequency $f_0$ is defined as the frequency at which the PSD has dropped to one-half its value at $f = 0$, we obtain

$$PSD(f) = \frac{k_BT}{\gamma\pi^2}\frac{1}{f^2 + f_0^2} \tag{2.9}$$

The prefactor here may for convenience be combined into a single fitting parameter $A = \frac{k_BT}{\gamma\pi^2}$, while $B = f_0^2$, leaving the function

$$PSD(f) = \frac{A}{f^2 + B} \tag{2.10}$$

to be fitted to the Fourier transformed data. From Equation 2.8 we obtain the trap stiffness from the fitted parameter $B$:

$$k = 2\pi\gamma\sqrt{B} \tag{2.11}$$

A Lorentzian curve, with the form [129]

$$L(x) = \frac{1}{\pi}\cdot\frac{\frac{1}{2}\Gamma}{(x - x_0)^2 + \left(\frac{1}{2}\Gamma\right)^2} \tag{2.12}$$

is then fitted and the $A$ and $B$ values found. From these, the stiffness of the trap is known.

In principle, the trap should be symmetric in $x$ and $y$, but in practice this is rarely the case. It is therefore required to perform this routine for both $x$ and $y$ data and subsequently work with separate force constants in each direction.

Having found the trap stiffness it is now necessary to obtain the displacement of the bead from the centre of the trap in order to find the force applied. Fortunately, this may be done precisely and reliably with the QPD. Observation of a bead trapped with no DNA tether gives an equilibrium voltage produced by the bead at the trap centre. This is normalised through division by $V_{sum}$ in order to eliminate effects deriving from differences in bead height relative to the trap focus [129].

QPD data which has also been normalised through division by the sum voltage may then be trivially subtracted from this equilibrium position to give the voltage difference at that time point:

$$\triangle V_i = \frac{V_{i,equil}}{V_{sum,equil}} - \frac{V_i}{V_{sum}} \tag{2.13}$$

where $i$ is one of $x$ or $y$.

Then $\Delta V_x$ to $x$ and $\Delta V_y$ to $y$ calibration must be generated experimentally to convert this difference in voltage to a meaningful position. There is a theoretical equation governing the QPD's response to bead displacement (here given for the $x$ axis; an identical equation describes $y$) [78]:

$$\begin{aligned} \frac{V_x}{V_{sum}} &\simeq \frac{16k'\alpha}{\omega_0^2\sqrt{\pi}} e^{-2\left(\frac{x}{\omega_0}\right)^2} \int_0^{\frac{x}{\omega_0}} e^{-\beta^2} d\beta \\ &\simeq \frac{16k'\alpha}{\omega_0^2\sqrt{\pi}} e^{-2\left(\frac{x}{\omega_0}\right)^2} \operatorname{erf}\left(\frac{x}{\omega_0}\right) \end{aligned}$$

which has a form when plotted as seen in Figure 2.17.

However, this theoretical curve should not be used on its own. Instead, for each microscope and after every change to the optical path, a calibration curve should be generated experimentally. This is done in principle by moving a bead fixed on the surface through the centre of the trap and in practice by performing a 2D scan using an immobilised bead, as a 2D raster ensures that the bead will pass through the centre of the trap. The central line is the line which maximises the QPD response in the direction of the scan (i.e. if the bead moves from top to bottom we would wish to maximise $V_y$) while minimising the response in the orthogonal direction (in the example of scanning through $y$, the $V_x$ response should be as small as possible). From knowledge of the movement of the stage, the mapping from voltage readout to displacement is then known.

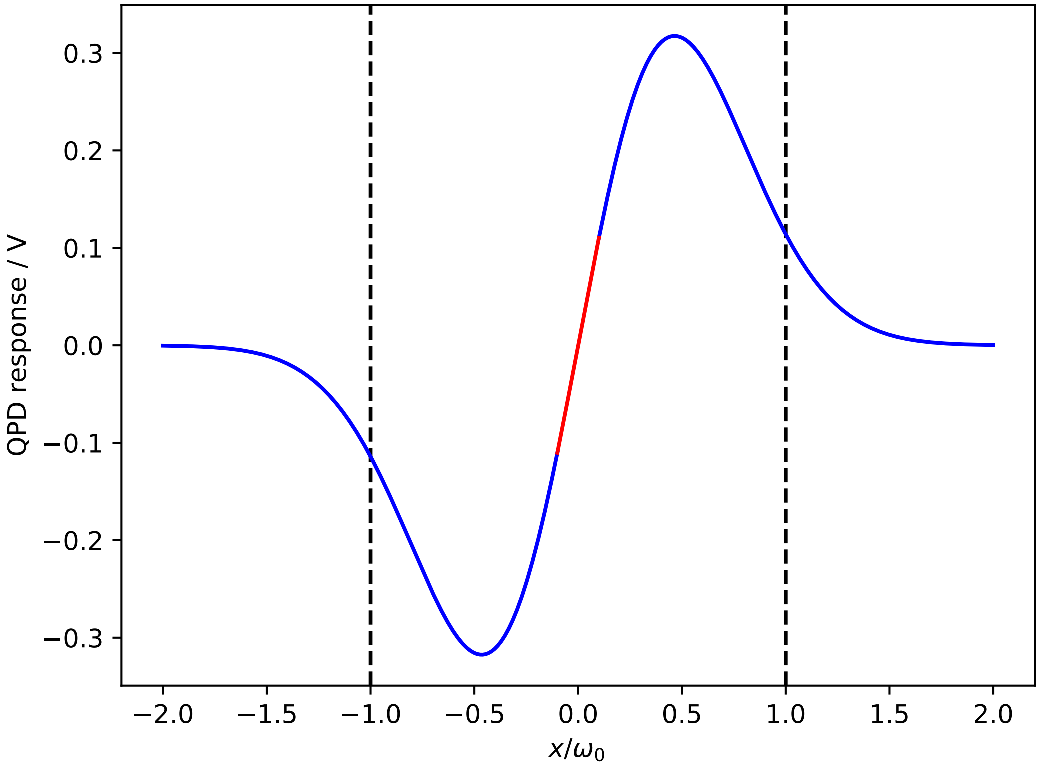


Figure 2.17: Theoretical QPD response to an object scanning through the centre of the QPD with linear region shown in red. $\pm\omega_0$, the 1/e width of the trapping laser's beam waist, is marked as dashed lines

However, as can be seen in Figure 2.17, the mapping is not one-to-one – one voltage gives multiple possible displacements. To eliminate this and to simplify the force calculations, we endeavour always to work in the central linear region of the trap. Here, the voltage-displacement mapping is single-valued, and the force applied by the trap remains Hookean.

Before any data was obtained, the QPD response was characterised. Significant deviation of this response from the theoretical curve above would suggest malfunction or misalignment.

To perform the characterisation, 3 $\mu$m magnetic beads were immobilised on the surface of a nitrocellulose treated cover slip following the same protocol as for the antiDIG beads. A bead was then chosen and a 2D raster scan was performed with the bead moving in both the $x$ and $y$ direction. The individual scans which minimised the response in one direction whilst maximising the response in the other were chosen and plotted.

The geometrical setup of the experimental equipment is such that it is only possible to use the $y$ axis as the axis of rotation for the magnetic bead. For this reason tethers are only formed with the DNA along the $y$ axis and Figure 2.18 shows only the QPD response along the $y$ axis.

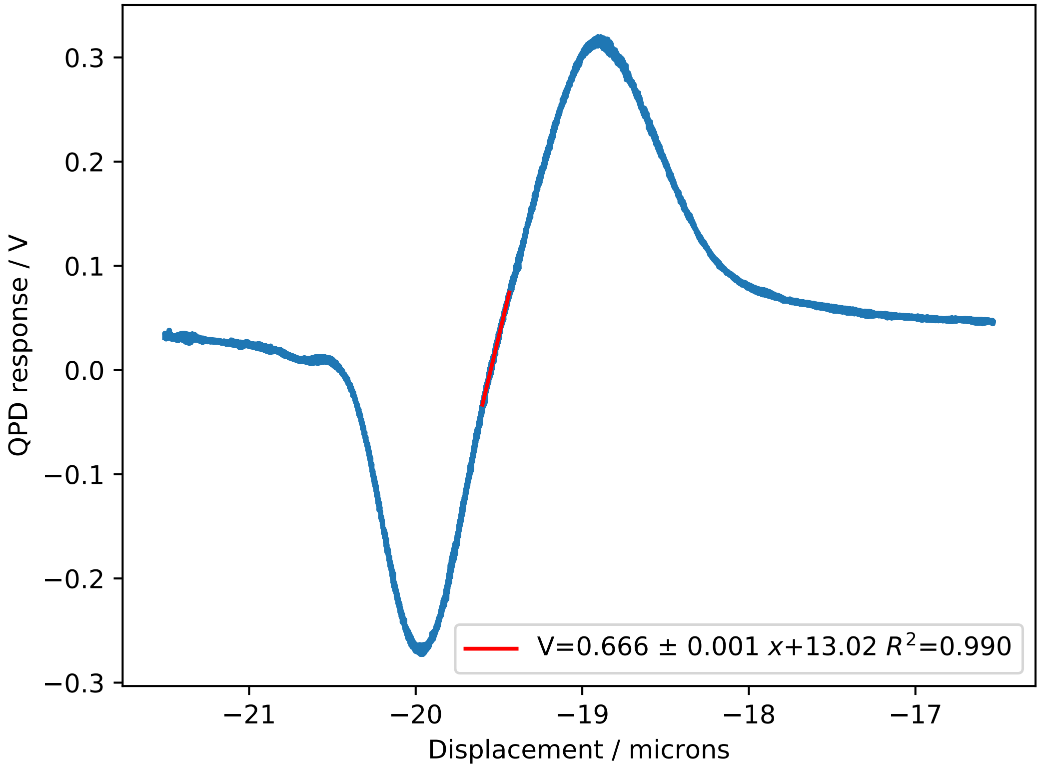


Figure 2.18: QPD response as a bead is scanned along the $y$ axis in the centre of the trap. The central region has been fitted to a straight line using the `curve_fit` routine in SciPy.

In Figure 2.18, it is clear that the experimental response is very similar to the theoretical curve in Figure 2.4. There is an approximately linear region in the centre of the response curve, which was fitted using SciPy's `curve_fit` routine. The sensitivity is given by the reciprocal of the gradient of the fitted line.

### 2.4.4 Quantifying Trap Stiffness

The stiffness of an optical trap is dependent on how far away from the cover slip surface the trap is positioned. For a trap too far into solution, spherical aberration introduced by the objective lens degrades the beam profile and weakens the trap. Conversely, a trap positioned close to the cover slip surface will find that strength varies due to local viscosity, an effect known as Faxen's law which operates over a depth around that of one bead radius. Faxen's law is given by [136]:

$$\beta = \frac{6\pi\eta a}{1 - \frac{9}{16}\left(\frac{a}{h}\right) + \frac{1}{8}\left(\frac{a}{h}\right)^3 - \frac{45}{256}\left(\frac{a}{h}\right)^4 - \frac{1}{16}\left(\frac{a}{h}\right)^5} \tag{2.14}$$

For an optical trap such as the one used in this project, the expectation would therefore be for trap stiffness to rise as the trap moves up from the cover slip surface and then to fall as spherical aberration degrades the trap quality.

This was tested by selecting magnetic beads (N>10) and trapping them with the optical trap. The bead was then brought down using the nanostage until it was touching the cover slip surface. It was clearly visible when this occurred due to the violent QPD response upon touching. The nanostage was adjusted such that the bead was just touching the cover slip, and QPD data was collected for 5 s. The nanostage was then moved such that the bead was 500 nm above the cover slip, and the bead was then allowed 4 s to reach an equilibrium position. Once again 5 s of QPD data was taken to determine stiffness. This procedure was repeated up to a height of 24 $\mu$m.

The collected data was analysed with two different means. The power spectrum method as described was used, as was the equipartition method. Briefly, the equipartition method takes the root mean squared displacement of the bead and relates that to a notional spring, the stiffness of which for the $x$ axis is given by

$$k_x = \frac{k_B T}{< \sigma \frac{V_{LR,x}}{V_{SUM}} >} \tag{2.15}$$

The equipartition method suffers some drawbacks however. Chief among these is the fact that low frequency mechanical noise can creep in and influence the stiffness – for example an experimenter coughing, touching the optical table, etc.

The power spectrum method is performed after removing low frequency contributions – the frequencies below 10 Hz were discarded when fitting the data obtained as part of this project. However, the fits can occasionally fail to converge and thus give poor quality answers. To remove this possibility, when the data collected for this experiment was fitted, each value was manually inspected and discarded if the fit did not give a reasonable estimate for the corner frequency.

For both equipartition and power spectrum data, the root mean square stiffness was found, and the averages and standard errors of the RMS stiffnesses were found and plotted.

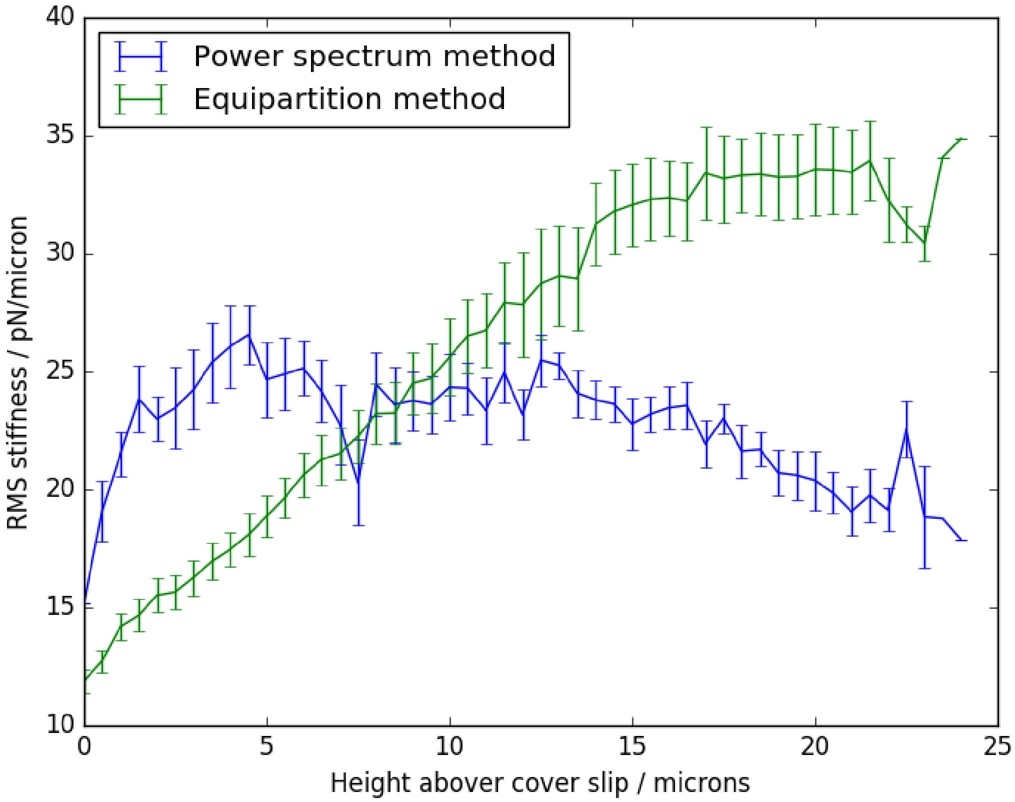


Figure 2.19: Trap stiffness as a function of height found with both the power spectrum method and the equipartition method. For all points, N>10. Bars: SEM

The results of this experiment can be seen in Figure 2.19. In the power spectrum data, as expected there is an initial increase in the stiffness followed by a gradual reduction. However, the equipartition method predicts a monotonic increase in RMS stiffness. This is likely due to mechanical noise, and the power spectrum method should be heeded as it is more reliable and qualitatively fits the expected behaviour.

### 2.4.5 Magnetic Tweezing

Many schema exist to manipulate particles with magnets, but this thesis will focus only on the method used for this work: application of torque with two pairs of Helmholtz coils, each pair controlled with an alternating current.

#### 2.4.5.1 Applying torque with Helmholtz coils

In order to apply only torque to a system, it is necessary to avoid applying force as well. Given that

$$\vec{F} = \nabla(\vec{m} \cdot \vec{B}) \tag{2.16}$$

and assuming that the value and orientation of $\vec{m}$ is constant, the only force which could act is due to an irregular magnetic field. Conversely, assuming a rotating uniform field, it is clear that the force acts only in an angular sense, and has a minimum when $\vec{m}$ and $\vec{B}$ are aligned. Thus in a uniform field, over time the magnetic dipole moment of the magnetic beads will align with the orientation of the externally applied magnetic field, without any additional spatial force being transduced. Viscous drag working on the surface of the bead means it takes time for the bead to respond to a new magnetic field, and so there is an upper limit to the rotational speed that can be achieved. However, previous work characterising these tweezers [111] demonstrated full rotation up to 7 Hz, so any rotation speed below this is acceptable.

To produce the necessary magnetic field, pairs of Helmholtz coils are used. These produce an extremely uniform magnetic field between the coils, and specifically along the centre axis. The magnetic field due to a current element is given by the Biot-Savart law

$$d\vec{B} = \frac{\mu_0 I}{4\pi r^2} d\vec{L} \times \hat{r} \tag{2.17}$$

where $d\vec{B}$ is the element of magnetic field, $\hat{r}$ is the unit vector from the current element to the point at which the field is to be evaluated, $d\vec{L}$ is the length element of the wire carrying current $I$, and the constants have their usual meanings. We consider a loop of wire with radius $R$, and evaluate a magnetic field a distance $z$ from its centre, along its central axis. The distance from the point to the wire loop is therefore $r = \sqrt{z^2 + R^2}$.

By symmetry, for a circular loop of wire we may neglect the vector formulation as the contributions in $x$ and $y$ cancel, so that

$$dB = \frac{\mu_0 I}{4\pi r^2} sin(\theta) dL \tag{2.18}$$

where $\theta$ is the angle between the $z$ axis and the vector between the location to evaluate and the current element considered, i.e.

$$sin(\theta) = \frac{R}{r} \tag{2.19}$$

Thus

$$dB = \frac{\mu_0 I R}{4\pi r^3} dL \tag{2.20}$$

or alternatively

$$dB = \frac{\mu_0 I R}{4\pi \left(R^2 + z^2\right)^{\frac{3}{2}}} dL \tag{2.21}$$

For our closed circle,

$$\begin{aligned} B &= \frac{\mu_0 I R}{4\pi \left(R^2 + z^2\right)^{\frac{3}{2}}} \oint dL && (2.22) \\ &= \frac{\mu_0 I R \times 2\pi R}{4\pi \left(R^2 + z^2\right)^{\frac{3}{2}}} && (2.23) \\ &= \frac{\mu_0 I R^2}{2 \left(R^2 + z^2\right)^{\frac{3}{2}}} && (2.24) \end{aligned}$$

For one Helmholtz coil of $n$ turns, the magnetic field at $z$ along its central axis is then

$$B = \frac{n \mu_0 I R^2}{2 \left(R^2 + z^2\right)^{\frac{3}{2}}} \tag{2.25}$$

The angular stiffness of the magnetic tweezers as designed and built by Dr Zhou was measured to be $3500\,\mathrm{pN \cdot nm \cdot rad^{-1}}$ at $\mathrm{B} = 4.8\,\mathrm{mT}$ for a $3\,\mu\mathrm{m}$ magnetic bead [111].

## 2.5 Final design

### 2.5.1 Coupling the fluorescence and trapping paths

The individual optical paths which make up the laser trap and the excitation path must be joined to make one cohesive machine. The optical table is organised so that each section of the microscope has its own dedicated footprint. To join them, a dichroic mirror is positioned at the back aperture of the microscope. The excitation light is reflected into the microscope, while the infrared (IR) beam strikes the dichroic mirror from behind and is transmitted through. This brings with it some practical complications. Primarily, the infrared beam is invisible and so alignment is generally performed with an IR viewer or laser card which fluoresces under excitation by near-IR light. However, the high number of optical components at the back of the microscope limits space to use this equipment.

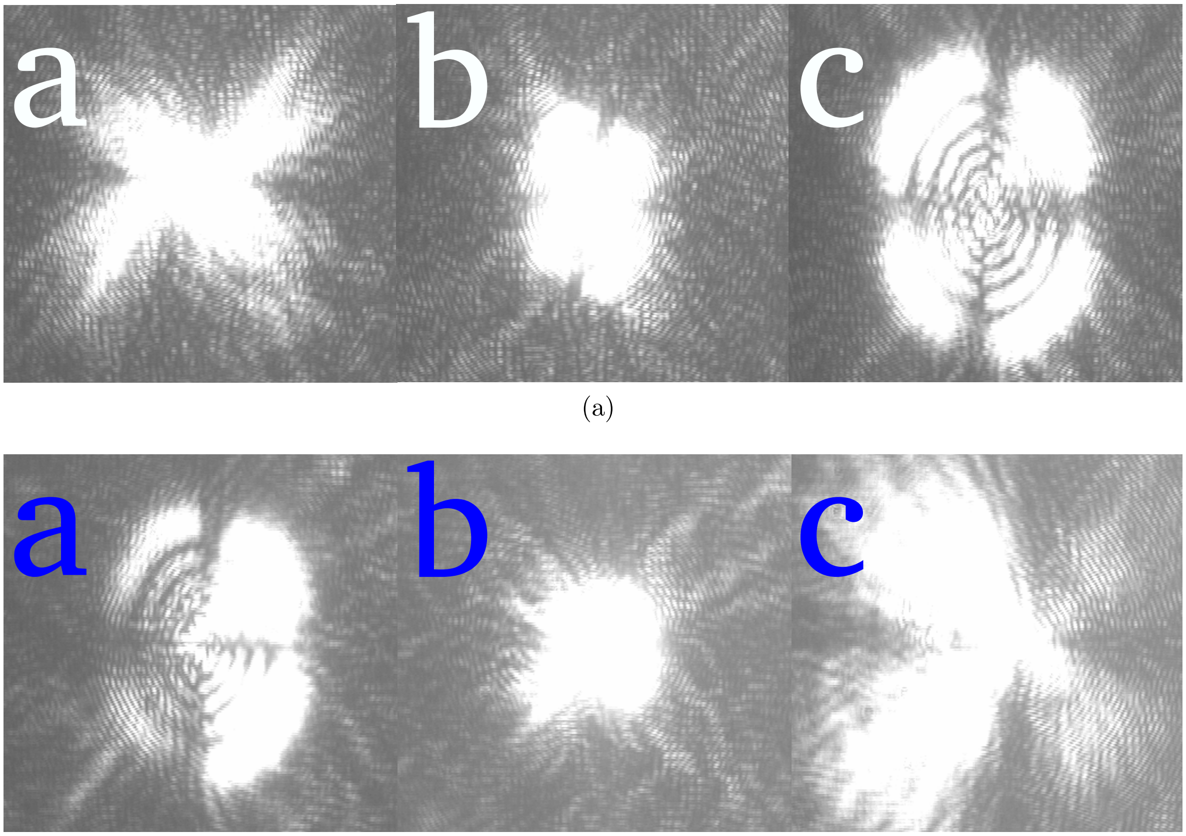
### 2.5.2 Aligning the optical trap

As described previously, the filter set in the Andor TuCam removes the IR light, making the beam impossible to see using the Andor cameras in this setup. Further, the dichroic mirror which is used for imaging YOYO-1 excludes IR as well. A second dichroic mirror is therefore used, a Semrock Brightline FITC-2024B-000, which allows approximately half of the trapping light through. The light is then taken out of the usual imaging path by way of a flip mirror and is passed to a Watec camera which is devoid of filters, allowing direct imaging of IR light by way of a custom LabVIEW program.

To align the trap, a water sample is loaded into the microscope and the surface of the cover slip found. The reflected IR from the cover slip surface produces a characteristic polarisation-dependent interference pattern, the shape of which depends on the height of the objective, as can be seen in Figure 2.20. When the laser is properly aligned, i.e. the IR light is incident perpendicular to the cover slip surface, there should be at different heights circular diffraction rings and a four lobed pattern, both symmetrical. By moving the focus of the microscope it is possible to see if the interference patterns are symmetrical across a range of heights. The trap being at an angle will manifest itself in the observed IR power going from one side to another as the objective moves up and down, as in Figure 2.20b. Clipping is demonstrated by part of the light being absent and the rest distorted.

In order to properly align the trap, a pair of mirrors is selected and one moved to make the pattern symmetrical, then the height is changed and the process repeated. This is done until symmetry across the range of heights is achieved, and in this iterative way it is possible to find the correct position for each mirror. To walk the beam – i.e. change its position – the second mirror is moved and the alignment performed again.

The Watec camera is also used for colocalising the excitation laser and the trapping laser. Each laser is turned on separately and their beams walked such that they are incident on the same area in the field of view. Though this is performed with a dichroic mirror different from the one used to image, if the beams are in the same place with one dichroic they will be with another, so it is not necessary to align with the imaging dichroic in place.



(a)

(b)

Figure 2.20: Images of nIR light reflected from a water sample.
a) The interference pattern with height when the nIR beam travels vertically. In all three it can be seen that the interference pattern is symmetrical
b) The behaviour changing with height with a non-vertical beam. The power moves from the right hand to the left hand side of the interference pattern indicating that the beam is incident at an angle to the cover slip.

### 2.5.3 Helmholtz coils and the custom slide holder

The pairs of Helmholtz coils mounted on the microscope chassis are subject to some design constraints. Firstly, the slide holder is large enough both the coils and the nanostage slide holder cannot be mounted simultaneously. Secondly, with the slide holder removed, there is a rectangular gap of a fixed size in the piezoelectric nanostage which all components need to fit in, and the oil-immersion condensing lens used is extremely wide. This limits space significantly and means that a commercial Helmholtz coil set is unlikely to be the solution.

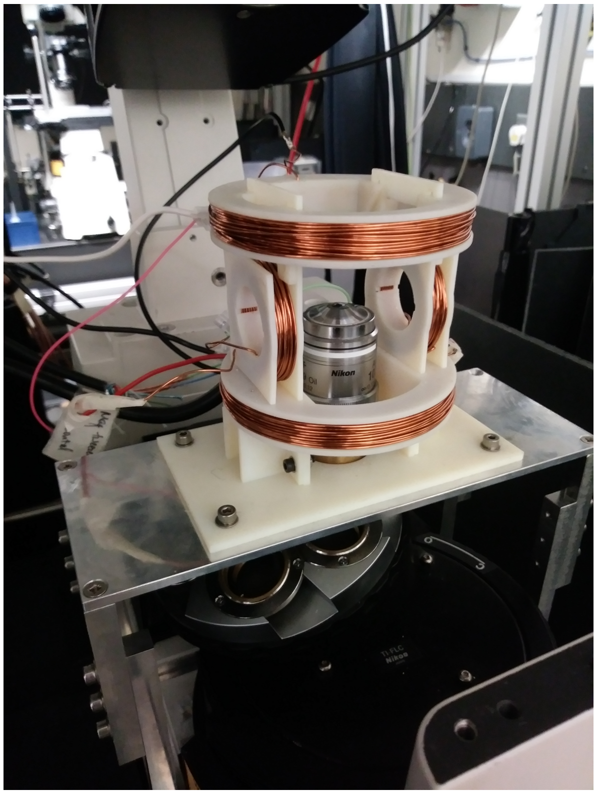


Figure 2.21: Photograph of the custom Helmholtz coils. It can be seen that the horizontal coils are considerably smaller than the vertical ones, hence needing higher current by a factor of 2.8 to achieve the same magnetic field.

Instead, two pairs of spools for the Helmholtz coils were designed and 3D printed by Dr Zhou as part of his PhD project [111]. They tessellate through butterfly joints and are mounted directly on to the microscope base, around the objective lens, a photograph of which is seen in Figure 2.21. This means that wherever the stage moves around the coils, the objective lens always images the sample at the centre of the coils and thus the area with the most uniform magnetic field. There only torque will be applied as discussed above. The coils themselves were wound with ceramic-insulated copper wire for good conductivity. The 3D printed material acts as a good heatsink, and the lack of an iron core in the coils further reduces heat generation. Because of this, overheating of the coils was not found to be a problem and therefore no cooling equipment was fitted. Due to the constraints on space, the range of movement of the stage with the Helmholtz coils in place is limited to approximately 1 cm × 1 cm.

Due to the rectangular shape of the hole and the large condensing lens, the vertical coils are larger than the horiziontal ones – 86mm diameter compared to 38.5mm (measured from the centre of the wires). The depth of the spool is also different leading to different winding – the smaller horizontal coils contain 60 turns of wire compared with

100. This means that the potential differences put across each pair of coils must be different by a factor of 2.8 to ensure that the magnetic field strength produced by both is the same in the centre of the sample.

To control the magnetic coils, a dual signal-generator/amplifier system is used. A waveform is generated to user specification by the LabVIEW control software, and this is transmitted to signal generators – one signal generator per coil pair. The generated sine waves are then sent to a pair of PC controlled linear power supplies (BOP 20-5M, Kepco Inc.), again one amplifier per coil pair. The coil pairs themselves are wired in series. The applied potential difference is monitored in real time through the measurement ports on the amplifiers, and written to a file via the DAQ card as detailed above. The potential difference applied across the coils is never greater in magnitude than 10 V. The coils have been shown to produce a trap with angular stiffness $1.1 \times 10^3$ pN·nm·$\mu$m$^{-1}$ at 15% power [111], and can produce a maximum magnetic field of 2 mT in the case of the larger pair of coils and 0.7 mT for the smaller pair [111].

## 2.6 Selecting magnetic beads

Optically trapping a bead is a matter of theoretical simplicity as demonstrated in sub-Section 2.4.1, but nature rarely allows us to see theoretical simplicity in practice and so it is here. The primary challenge when selecting a bead for this work is that magnetic materials and optical trapping are fundamentally incompatible. Metal present in a trapped bead absorbs IR and can therefore overheat the bead [137]. This absorption also increases the radiation pressure exerted on the bead, leading to a weaker trap in the best case or the bead exiting the trap altogether. If too much of the magnetic material is removed, however, the beads may become so weakly magnetised that they are not affected by the magnetic tweezers. It is also true that smaller beads are more easily trapped (with these tweezers, optically trapped 1 $\mu$m beads were found to have a trapping stiffness an order of magnitude higher than that of the magnetic beads) – but in this instance if we reduce the size of the beads, the effective concentration of ferrite particles increases and the undesirable properties outlined increase in severity.

The distribution of ferrous materials is also of concern. A ferrite centre with a polymer around it is an optically different system from a random coating of magnetic particles forming a shell on the surface of a bead, and the choice of distribution is important – in this work, no beads with a solid ferrite core were reliably optically trappable. Finally, any magnetic particle used for trapping must be functionalised with Streptavidin, Avidin, or Neutravidin for conjugation with the DNA. This may be done either by the manufacturer or in-house but to reduce the chances of error a commercial choice is preferable.

Several magnetic bead products were identified and trialled for these purposes, but only the SpheroTech Micromer-M streptavidin-functionalised beads were both optically and magnetically trappable, so were chosen for this project.

The beads are made of a styrene-maleic acid copolymer, are superparamagnetic, and are of diameter 3 $\mu$m with low variation in size as seen in Table 2.1. However, the distribution of $Fe_3O_4$ varies much more dramatically at between 9 and 14% by weight. Therefore the beads cannot be thought of as uniform, and so when performing any measurements with these beads it is necessary to characterise each bead individually. Smaller, non-magnetic beads would usually be characterised *en masse* and the population averages referred to but the high heterogeneity of the beads used here render that approach unacceptably inaccurate. The distribution of trapping stiffness, which demonstrates this, is seen in Figure 2.34 in Chapter 2.

| | |
|---|---|
| Diameter | $3\mu$m |
| Coefficient of variation of diameter | $< 5\%$ |
| Density | 1100 kgm$^{-3}$ |
| Magnetic material | $Fe_3O_4$ |
| Magnetic material content | 9-14% by weight |
| Iron content | 6-10% by weight |
| Magnetisation per bead | $8.4 \times 10^{-14}$ Am$^2$ |

Table 2.1: Physical properties of the chosen magnetic bead

## 2.7 Demonstrating trapping and rotating

Before performing any measurments, the core capabilities of the machine were demonstrated. Firstly, a magnetic bead was trapped optically and held at 1, 5, and 10 $\mu$m above the cover slip surface. The results of this can be seen in Figure 2.22. Near to the cover slip surface, the bead is found to be more out of focus due to high radiation pressure as seen in Section 2.4.1, but as the bead was taken deeper into solution the restoring force experienced by the bead had a greater influence and the bead came back into focus more. The trap stiffness with height is shown in Figure 2.19 in Chapter 2. Though this defocussing raises a problem, it shows that the optical trap operates correctly over a range of heights and can successfully trap magnetic beads.

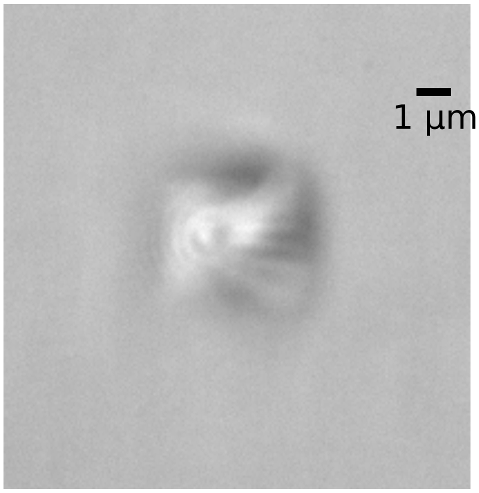


Figure 2.22: A trapped bead 1 $\mu$m above the cover slip. At this experimental working height, the bead is significantly out of focus. Being a small molecule, DNA would not be visible this far from the focal plane of the objective lens. It is therefore necessary to readjust the focus with additional optics in the emission path.

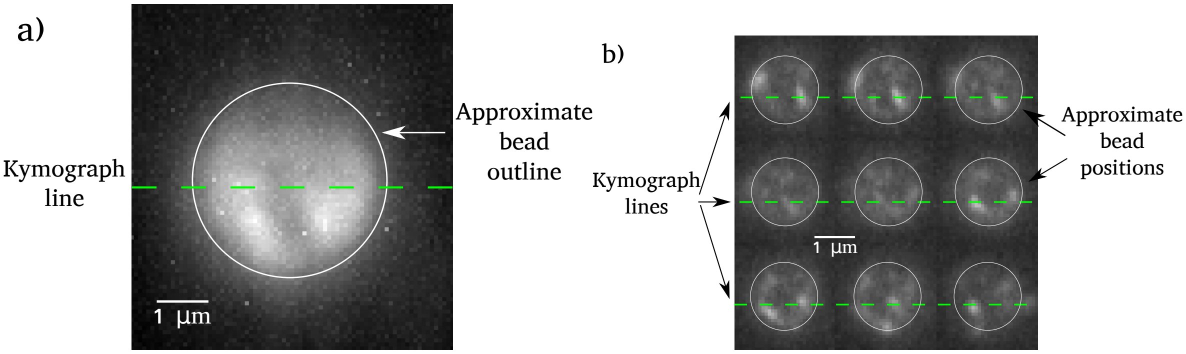


Figure 2.23: Fluorescence images of a trapped bead rotating due to the magnetic tweezers. a) An intensity average over 200 frames. The green circle indicates the approximate bead position. The brightness distribution compared to the montage images demonstrates that the bright spots must be moving. b) A montage of the first nine frames of the fluorescence movie. The bright spots can be seen to move from one frame to another though the direction and rate of movement is unclear. In both a) and b) the lines used for taking the subsequent kymograph are indicated.

The magnetic tweezers were next to be tested. The magnetic beads were fluorescently tagged with biotinylated beads which fluoresce over the same range as YOYO-1 and are therefore suitable to use as fluorescent fiduciary markers. A bead was selected and trapped in the optical trap and the Helmholtz coils were set to rotate at 1 Hz. Imaging in fluorescence, movement of the fluorescent beads is easily seen, and is shown in Figure 2.23. However, in still images the sense of the movement is unclear. Therefore a kymograph was taken along the line indicated in Figure 2.22, and the result is shown in Figure 2.23. The slanted lines in the kymograph show that fluorescence moves from right to left through time, and these lines are repeated at regular intervals. The lines are all at the same angle relative to the time axis, which indicates that the bead is not simply rocking back and forth. Taken together, this indicates that the bead is rotating fully at a constant rate, and that the bead remains in the optical trap under this additional manipulation.

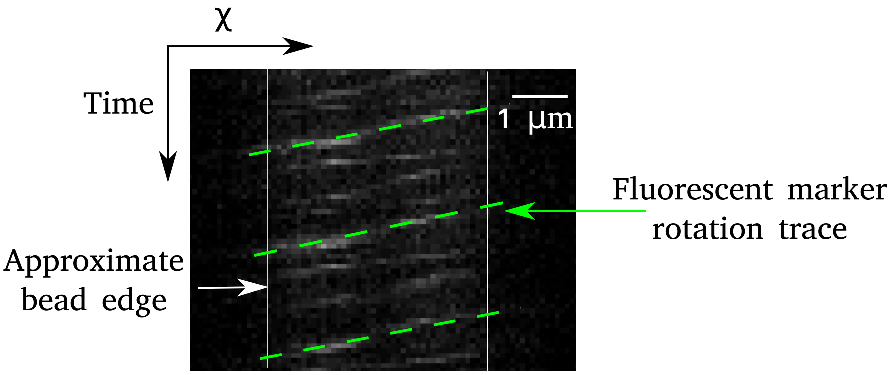


Figure 2.24: Kymograph of the optically trapped, rotating bead. The indicated slanted lines indicate fluorescent spots moving from right to left through time. The repetition of the slanted lines indicate that the movement is also repeated, and thus that full rotation is ocurring. The lines are consistently at the same angle, showing that the rotating is at a constant rate. The edges of the bead remain in the same place through time showing that under rotation the bead remains in the optical trap.

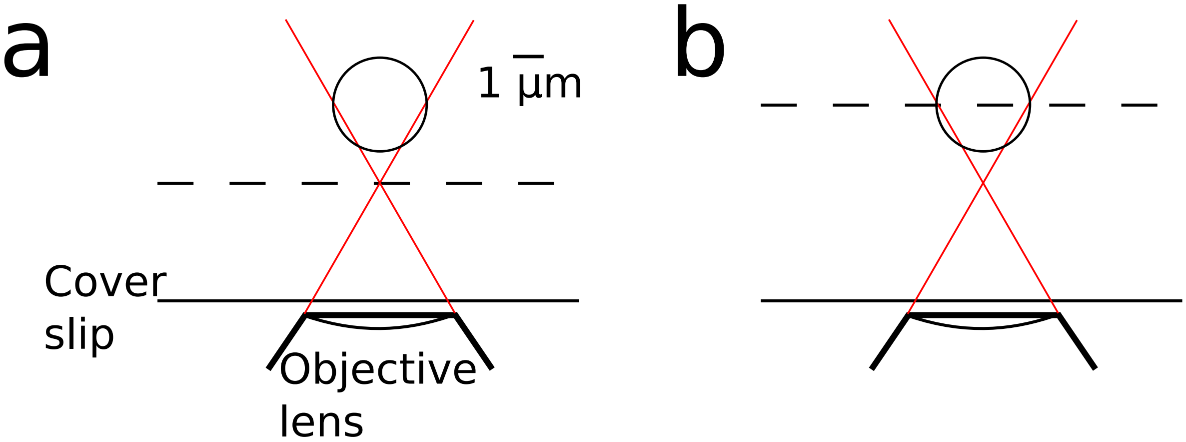


Figure 2.25: The shift in imaging plane due to the corrective module in the imaging path. The IR beam is denoted by the red lines, and the dashed lines indicate the imaging planes.

### 2.7.1 Correcting the focus

Although above it was seen that both primary functionalities of the combined tweezers worked as hoped, one problem was clear: with the bead centre around 2.5 $\mu$m above the cover slip (where the centre of the surface immobilised anti-digoxigenin bead will be, and where experiments will take place) the trapped bead is significantly out of focus. To correct this, the in focus plane was shifted by means of a corrective optical module inserted into the imaging path, the effect of which is seen in Figure 2.25.

To calculate the strength needed for this corrective module, the thin lens Equation is used:

$$\frac{1}{f_{corr}} = \frac{1}{f_{objective}} - \frac{1}{\delta z} \tag{2.26}$$

where $f_{corr}$ is the focal length of the correction module, $f_{objective}$ is the focal length of the objective lens, and $\delta z$ is the distance that the focal plane needs to move, i.e. the height that the forward radiation pressure pushes the trapped bead. Experimentally it was determined that $\delta z \simeq 2\,\mu$m, and it is known that $f_{objective} = 2$ mm. Thus

$$f_{corr} = \left(\frac{1}{2} - \frac{1}{2 \times 10^{-3}}\right)^{-1} \text{mm} \simeq -1.9\,\text{m} \tag{2.27}$$

From the lensmaker's equation:

$$\frac{1}{f_{corr}} = \frac{1}{f_1} + \frac{1}{f_2} - \frac{d}{f_1 f_2} \tag{2.28}$$

where $f_1$ and $f_2$ are the focal lengths of the two lenses which make up the correction module and $d$ is the distance that separates them. As space in the imaging path is at a premium, the interlens separation was the parameter around which the others were optimised. Both lenses were to be mounted on a ThorLabs lens cage on a breadboard, which was affixed to the optical table with magnetic clamps for ease of removal. The total space for this was around 100 mm, so Equation 2.28 was solved by limiting $d$ to around 50 mm. It was found that $d = 48$ mm, $f_1 = -75$ mm and $f_2 = 50$ mm satisfied the equation and fit in the space well. These are also standard lens focal lengths, and allowed some movement of the lenses on the cage to give flexibility to cope with slightly different values of $\delta z$.

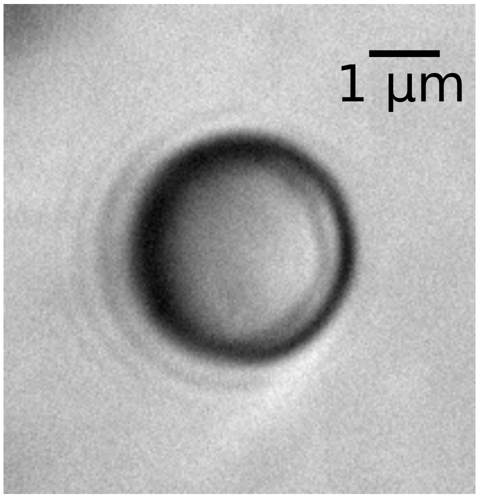


Figure 2.26: The effect of the corrective module on the appearance of a bead optically trapped 1 $\mu$m above the cover slip. The bead now appears in focus, indicating that the focal plane has been shifted correctly. Horizontal DNA at the level of this bead should be visible using fluorescence microscopy.

The result of this corrective module can be seen in Figure 2.26. The bead is now in focus across a much wider range of heights, including the important height of 2.5 $\mu$m. It also magnifies the sample still further. Without the corrective module, the microscope captures images at 80 nm/pix but with the module in place, the resolution is 32.5 nm/pix.

### 2.7.2 Force clamping

Finally, code was written by Dr Zhou to create a force clamp using the nanostage and the information from the QPD. The inhomogeneity inherent in the system – due to ferrite content and distribution, bead roughness, beads not being perfectly spherical, or a DNA tether formed at the side of the bead instead of along the axis of rotation – mean that as they are rotated there is periodic noise introduced to the QPD reading which has a period commensurate with that of rotation [111]. To force clamp effectively this noise should be removed. Data is therefore collected over one rotation of the magnetic tweezers and an average found. As well as reducing the periodic noise due to rotation, this is a sufficiently long time to effectively remove Brownian noise from the data too.

This averaged QPD reading is then converted to a force as outlined previously. The calculated force is compared to the user-specified desired force, and the approximate distance between the current and target bead displacement is found. For example, if the desired force is 2 pN and the stiffness of the trap is 20 pN$\cdot\mu$m$^{-1}$, the bead should on average be 100 nm from the trap centre. If the bead is found to be on average 50 nm from the trap centre, the trap must be moved. The distance to move is calculated in the following way:

$$< \delta F > \quad = \quad F_{target} - < F_{measured} > \tag{2.29}$$

$$\delta x \quad = \quad \frac{< \delta F >}{S_{trap}} \tag{2.30}$$

where $S_{trap}$ is the measured sensitivity of the trap expressed in pN$\cdot\mu$m$^{-1}$.

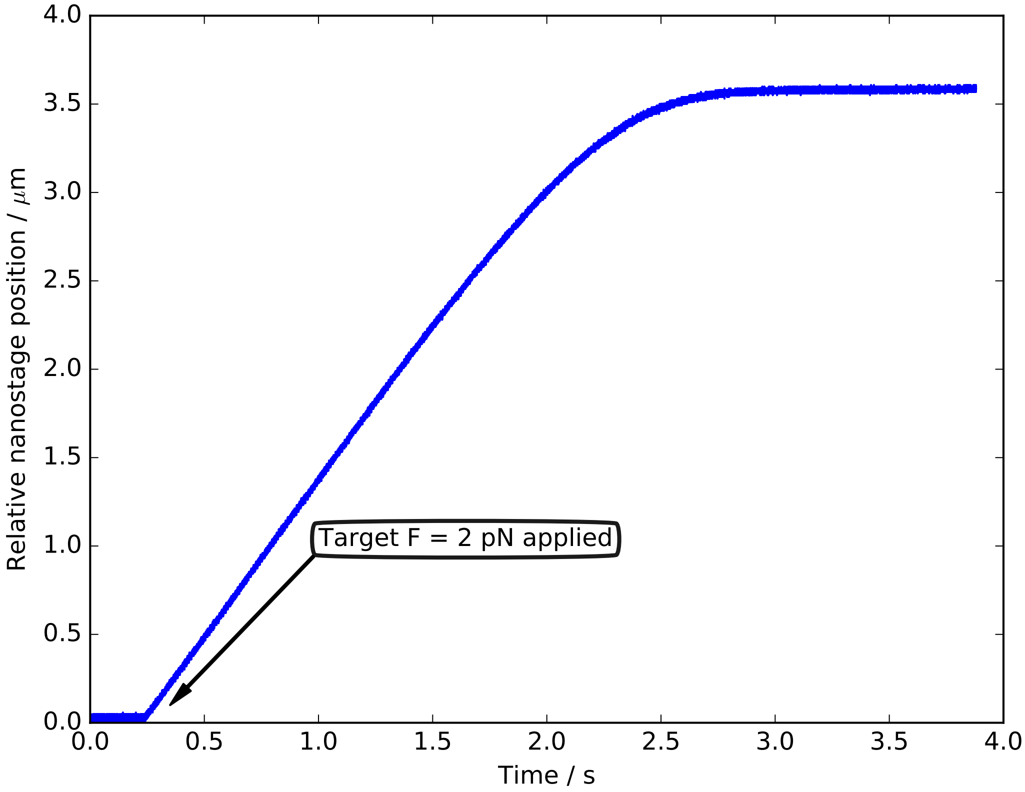


Figure 2.27: The response through time of the force clamp to sudden application of a 2 pN target force starting at $F = 0$. As would be predicted, the response is initially linear as the trap moves at its maximum speed. As it gets closer to applying the desired force, the response becomes nonlinear and the target is gradually reached over a period of seconds.

Though in principle it is possible to move the nanostage the exact distance required, this can become unstable, apply jerking motions to the sample, and will irretrievably break down if a jump is too far. Instead, the requisite distance is calculated and the nanostage moved by a fraction of that – usually half. This results in a smooth, stable, noise-resistant movement towards the desired force, which should decay exponentially as it moves half the distance each time.

Before force clamp experiments began in earnest, the response of the force clamp was assessed by sudden application of a target force of 2 pN when the clamp was in a position where the force was 0. As is seen in Figure 2.27, the force clamp begins moving immediately the target is applied, and reaches its target within a few seconds. The quick onset of movement and the smooth response suggest that the force clamp should be able to maintain a constant force in these experiments – the DNA should not produce leaps in force of above 2 pN.

## 2.8 Bead Assays

### 2.8.1 Surface Immobilisation of the antiDIG Beads

The beads chosen to act as anchors were 4-4.9 $\mu$m antidigoxigenin (antiDIG) coated polystyrene beads from SpheroTech (part number DIGP-40-2). These are larger than the 3 $\mu$m diameter magnetic beads for use with the optical trap, allowing a tether to be formed without the optically trapped bead touching the surface of the cover slip. The antiDIG beads also have the advantage of being extremely stable – when stored at 5°C a test batch lost only 5% of its activity over a period of 3 years [138].

The antiDIG beads need to be firmly attached to the surface of the cover slip in order to act as an anchor for the DNA tether. Previous studies have made use of conjugated DIG-BSA [139] to immobilise large beads. However in that previous work a larger bead was used and was sandwiched between the cover slip and microscope slide. As a result, the stress applied to the BSA-glass coupling was reduced. In the experimental geometry in this thesis, a deeper chamber was used and the flow through it therefore exerted a force on the immobilised beads, removing them from the surface. Ultimately, DIG-BSA was found not to be a strong enough attachment point for the purposes of this project.

Instead, a non-specific electrostatic immobilisation was used. Nitrocellulose has been used as an adhesive in microscopy for some time, especially for immobilising antibodies on a cover slip surface. To test its efficacy with the antiDIG beads, nitrocellulose (09817, Sigma-Aldrich) was diluted to a final concentration of 0.1% in amyl acetate (W504009, Sigma-Aldrich). Cover slips were prepared for functionalisation by plasma cleaning for 1 minute as described in [119] and were then submerged in the nitrocellulose solution. The treated cover slips were allowed to air dry and were then immediately formed into tunnel slides, which were made of two strips of double-sided Scotch brand tape placed approximately 5 mm apart on a microscope slide. The treated cover slip was then dropped on to the tape, forming a cuboidal chamber into which solutions could be introduced.

Since the antiDIG bead-nitrocellulose interaction is electrostatic in nature [140], it is important that the correct ionic strength solution is used to allow the beads to get close enough to the surface to form a strong bond. Phosphate-buffered saline (PBS) tablets (18912014, ThermoFisher Scientific) were dissolved to their advertised 1X concentration, consisting of 10 mM phosphate, 2.68 mM potassium chloride, and 140 mM sodium chloride [141]. To this buffer was added different concentrations of additional potassium chloride, specifically 100 mM, 200 mM, 300 mM, and 400 mM. The bottle of antiDIG beads were vortexed for 30 seconds to resuspend all beads, and 20 $\mu$L of the stock placed in each of 5 microcentrifuge tubes. The volumes were made up to 100 $\mu$L with one of each of the five buffers (PBS and the four PBS + KCl solutions). The beads were mixed by pipetting and all five tubes centrifuged in a bench-top centrifuge at maximum speed for 5 minutes to form a bead pellet. After centrifugation the supernatants were removed and the pellets resuspended in the buffer used to dilute the beads, to a final concentration of 60 $\mu$L. Dilutions of beads in five different ionic strength buffers were therefore obtained.

Four tunnel slides for each buffer were made, and into them introduced 15 $\mu$L of diluted beads. The slides were inverted and kept in a sealed humidity chamber. After 5, 15, 30, and 60 minutes one slide from each ionic strength buffer was taken and washed with 200 $\mu$L of the chosen buffer. The slide was then put on a bench-top microscope and imaged at 400× magnification with a field of view 500 $\mu$m × 500 $\mu$m. Five fields of view from each slide were chosen at random and the immobilised antiDIG beads present counted. The number of beads per field of view for each incubation time and condition were then averaged and a standard error was calculated from the number of beads in each of the five fields of view. One slide was performed for each salt condition and each incubation time, so the standard error in the mean refers to the standard error across the observed fields of view in one slide, not across multiple slides.

As can be seen in Figure 2.28, the higher ionic strength buffers outperformed the weaker ones. All slides showed a decrease in immobilised beads as time went on, this likely being due to weaker attachments breaking down over time. The PBS + 400 mM KCl buffer was the best performer, with consistently higher numbers of beads per field of view, and low standard error. Additionally, the number of beads per field of view plateaued earlier than for other solutions and so a 1 hour incubation would be sufficient to ensure poorly immobilised beads would not be present. Therefore PBS + 400 mM KCl was chosen to be the bead incubation buffer when preparing samples.

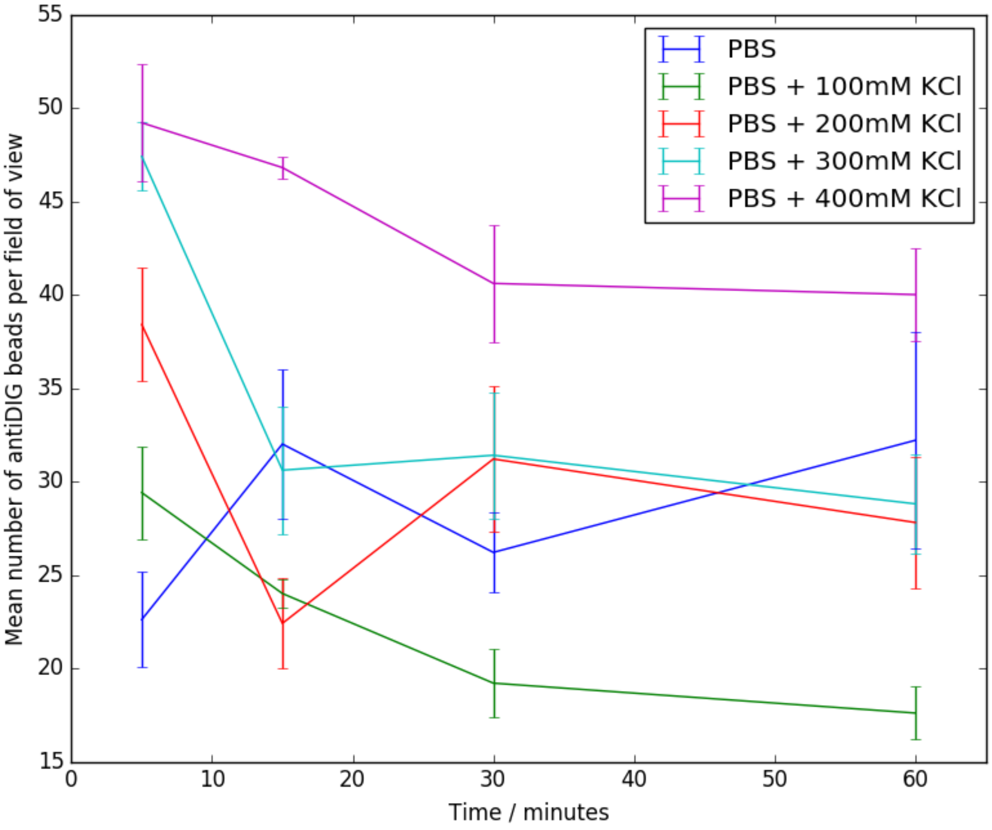


Figure 2.28: Mean number of immobilised 5 $\mu$m antiDIG beads per field of view on the benchtop microscope in different ionic conditions. 5 fields of view were observed for each data point. Bars: S.E.M.

Following the incubation, the surface of the cover slip was passivated by incubating the slide with 2 mg/mL BSA (A7638, Sigma-Aldrich) dissolved in PBS.

### 2.8.2 Conjugation of the DNA Construct

After bead immobilisation and the subsequent passivation, the slide was incubated with the DNA construct. With an average of around 40 beads per field of view on the bench-top microscope, and a field of view of 500 × 500 $\mu$m, there would therefore be an average of 17,600 beads per tunnel slide.

For incubation, the 4 kbp tether was diluted in PBS to a final concentration of 2 pM, while the 2.15 nM 15 kbp DNA construct was diluted in 40mM Tris-HCl +

10 mM MgCl2 to a final concentration of 8.6 pM. 20 $\mu$L of either tether was flowed into the tunnel slide to ensure the entire volume was the replaced by the diluted DNA construct. The slide was then inverted and incubated for 15 minutes in the case of the 15 kbp tether and 30 minutes for the 4 kbp construct. The correct incubation times were determined through fluorescence microscopy assays.

For the 15 kbp tether, 20 $\mu$L 17.2 pM DNA was mixed with 20 $\mu$L 1 $\mu$M YOYO-1 (Y3601, Invitrogen) and imaged with epifluorescence microscopy. The length of the DNA was such that the DNA molecules could be seen as defocussed strands coming off the beads. For some beads, the DNA tethers nonspecifically adsorbed to the surface and were thus held in an extended geometry. Representative frames of a tether in that orientation are shown in Figure 2.29. These images demonstrate that an incubation of 15 minutes is sufficient for DNA-bead complexes to form.

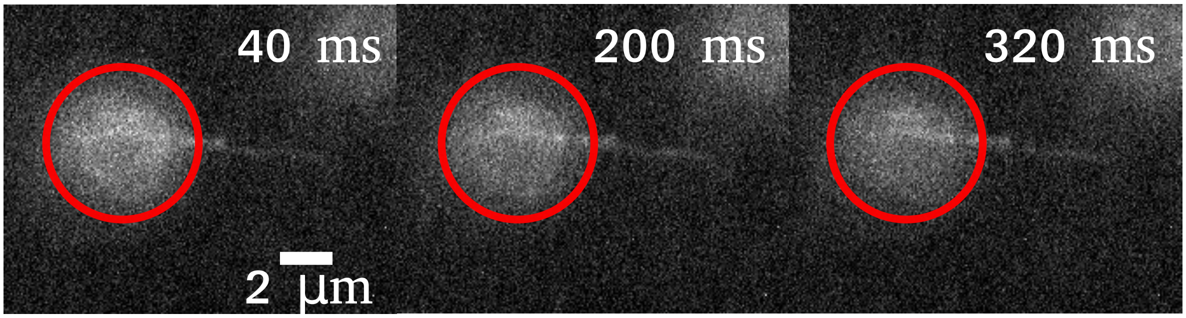


Figure 2.29: Subsequent frames from a fluorescence acquisition of the 15 kbp construct conjugated to an antiDIG bead. Exposure time 40 ms. DNA is visible as a bright line of fluorescence visible on the right hand side of the antiDIG bead

The 4 kbp construct has a significantly shorter contour length than that of the 15 kbp DNA, and as a result the DNA could not be observed directly. Instead, the 4 kbp tether was diluted to 4 nM and incubated at a ratio of 1:1 with 250 nM YOYO-1. Fluorescence images of the antiDIG beads were captured every ten minutes for an hour. The first 100 frames of the acquisition were averaged and the total intensity of the antiDIG bead taken. Five beads were imaged at each time and the means and standard errors plotted. The result of this can be seen in Figure 2.30. Clearly, in this case 15 minutes would not be sufficient as the fluorescence is on the same scale as the control sample, while the leap in fluorescence at 30 minutes indicates binding of DNA has succeeded.

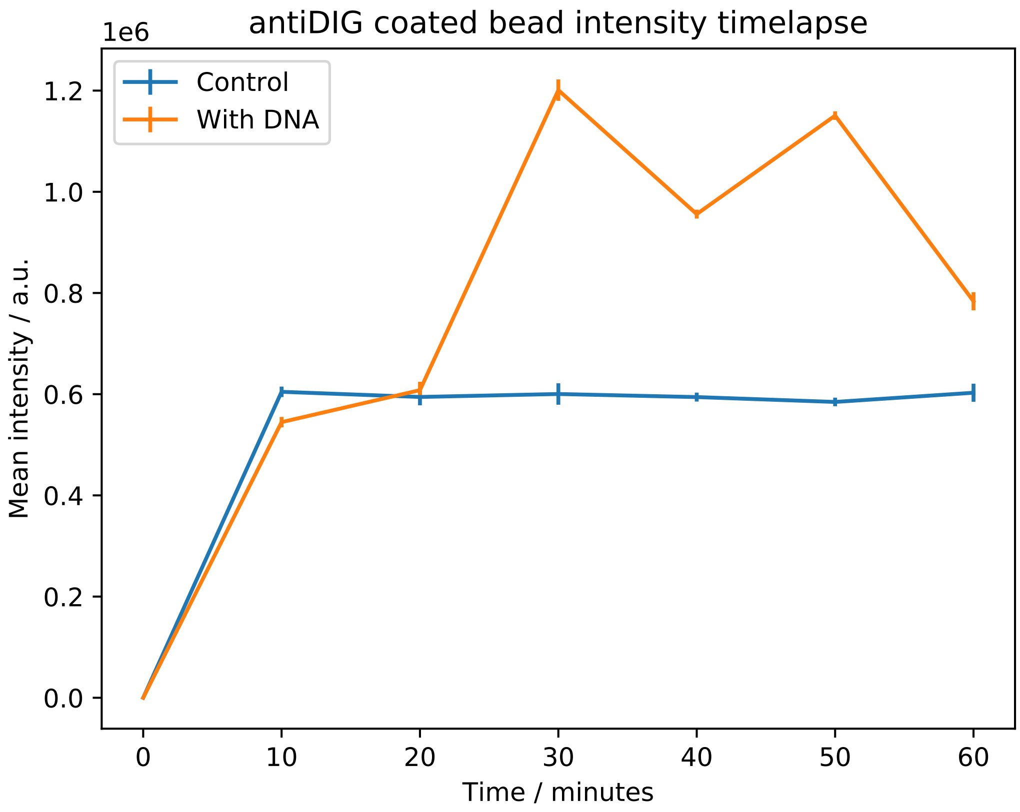


Figure 2.30: Total fluorescence intensity for antiDIG beads through time. Control data was taken using the same concentration of YOYO-1 with no DNA present. One slide was examined over time to gather data for each point, and the number of beads imaged at each time was 5. Error bars represent the standard error in the mean for these 5 beads calculated from the mean and standard deviation of the bead intensities with N=5.

A tunnel slide has approximate dimensions 5 × 22 × 0.1 mm, and hence a volume of approximately 10 $\mu$L. The approximate total number of 15 kbp DNA construct molecules incubated in the tunnel slide was $86 \times 10^{-18} N_A = 517.72 \times 10^5$ where $N_A$ is Avogadro's constant with the value $6.02 \times 10^{23}$. The approximate number of 4 kbp tethers was higher at $2408 \times 10^5$ as tether formation was found to be more difficult with this construct. On average this implies that for each antiDIG bead there were either 3000 15 kbp DNA constructs or 14,000 4 kbp DNA constructs. However, as the DNA molecules are dependent on Brownian motion for movement, in practice fewer than that would be expected to interact with each bead and fewer still to conjugate properly. During experiments, the above concentrations were qualitatively found to be sufficient to form DNA tethers regularly but low enough that multiple tether formation was a rarity.

### 2.8.3 Tether Formation

After 15 minutes the excess DNA from the incubation was washed out with 200 $\mu$L of a wash buffer consisting of 40 mM Tris-HCl pH 7.5 and 10 mM $MgCl_2$. 20 $\mu$L of imaging buffer were then introduced into the tunnel slide. The imaging buffer is based on T4 DNA Ligase buffer, and contains the free radical scavenging system as well as T4 DNA ligase to heal single-stranded nicks when they occur. The composition of the T4 ligase buffer is as follows: 50 mM Tris-HCl (pH 7.5), 10 mM $MgCl_2$, 10 mM DTT, 1 mM ATP [125]. The final buffer also contains 2 mg/mL Trolox , 15% glucose, 1 mg/mL glucose oxidase (G7141, Sigma-Aldrich), 0.04 mg/mL catalase (02071, Sigma-Aldrich), 0.2 mg/mL BSA, and magnetic beads at a final concentration of 1:999 from stock.

To form a tether, a properly immobilised antiDIG bead was selected and a nearby freely-floating magnetic bead was optically trapped. The corrective module was put in place in the imaging path and the magnetic bead was brought towards the cover slip until both the antiDIG bead and the magnetic bead were in focus. A measurement of the magnetic bead's Brownian motion was then taken using the QPD for determination of trap stiffness. The two beads were then brought to within 500 nm of each other and left for 2 minutes for tethers to form.

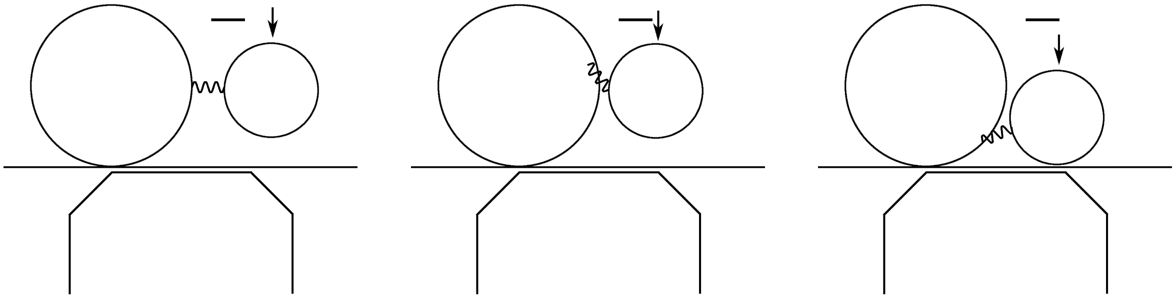

(a) (b) (c)

Figure 2.31: Possible tethering geometries for the 1 $\mu$m DNA tether, drawn to scale. Bars 1 $\mu$m. The orientation of the optical trap is indicated by the arrow. a) The ideal tethering geometry. The maximal amount of the contour length is visible horizontally to the field of view. b) A suboptimal tethering geometry. Only some of the contour length is visible, and that at an angle to the horizontal. c) The worst possible tethering geometry. The entire contour length is obscured by the beads

To evaluate if the tether had successfully formed, the beads were separated slowly using the nanostage. If a single tether had successfully formed, the bead would begin to be pulled from the trap at separations close to the contour length of the construct. If this did not occur, or the bead was pulled from the trap at a close separation, tether formation was judged to have failed or multiple tethers to have formed, and in either case a new pair of beads was sought. If a single DNA molecule appeared qualitatively to be present, this was quantified by oscillating the trapped magnetic bead while recording the displacement from the trap using the QPD. The force extension curve generated by this method was later plotted, and fitted to the wormlike chain model [33].

However, both tethers were not equally suited to forming tethers for use in this kind of microscopy based assay. Figure 2.31 shows possible binding geometries for the 4 kbp tether. It can be seen that it is most likely for a tether to form at an angle – not perfectly horizontally from bead to bead. For the longer tether, this is not in practice a problem. The contour length sufficient for the beads to be comfortably separated by several micrometres. This is not the case for the 4 kbp tether, which has a maximum bead-bead displacement of 1 $\mu$m. This is a problem for two reasons. First, the beads not being separated makes imaging them with fluorescence microscopy impossible – the signal will be hidden by the beads. Second, and more importantly, when the beads are close together, the IR beam which both traps and measures the magnetic bead is encroached upon by the immobilised antiDIG bead. This introduces noise and errors into the QPD data and makes extracting meaningful information difficult if not impossible. All tethering assays performed with the 4 kbp construct went similarly. The beads were incubated close together and slowly drawn apart using the nanostage. However, before there was any open space between the two beads, the magnetic bead was pulled from the trap and springs back close to the antiDIG bead. The poor quality QPD data and low bead separation with this tether meant that it was unlikely that biological insight could be drawn from the shorter tether, and it was therefore rejected. All the data which follows in this thesis was collected using the 15 kbp tether under conditions described above.

## 2.9 Removing Excess Dye

During initial fluorescence assays it was noticed qualitatively that the antiDIG beads were extremely bright and in fact saturating the camera. The experiments were being performed in an excess of YOYO-1, and it was hypothesised that the excess unbound dyes were immobilised on the antiDIG surface and contributing to the fluorescence. To test this, the excess YOYO-1 was removed using a commercial size-exclusion kit (7326250, Bio-Rad Laboratories) according to the manufacturer's instructions. The

DNA and YOYO-1 in the same concentrations as sub-Section 2.8.2 were incubated and eluted in a buffer consisting of 40 mM Tris-HCl (pH 7.5) + 10 mM $MgCl_2$. The purified and unpurified DNA-YOYO-1 complexes were incubated with the surface-immobilised antiDIG beads for 15 minutes at room temperature in a humidity chamber, and imaged using epifluorescence microscopy as described in Chapter 2.

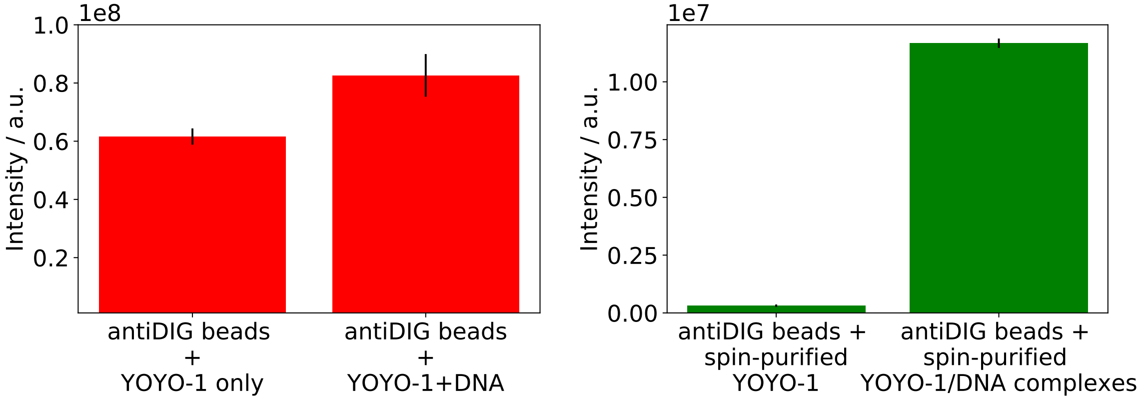


Figure 2.32: Bar charts showing a) integrated fluorescence intensity of beads in the presence of YOYO-1 and YOYO-1/DNA complexes, and b) intensity of beads with YOYO-1 and YOYO-1/DNA complexes after size exclusion column purification. In a) the YOYO-1 without DNA present is approximately 75% as bright as those with both YOYO-1 and DNA present, despite the fluorescence boost experienced by YOYO-1 on binding to DNA. In b) after purification with the size exclusion column, the beads conjugated to YOYO-1/DNA complexes are approximately 30× brighter than the beads in the presence of post-purification free YOYO-1, demonstrating that the size exclusion column works as expected and that the DNA remains fluorescent after purification. For each bar, 20 beads were sampled, and the error bars represent the standard error in the mean of these beads with N=20.

100 frames with an exposure time of 40 ms were taken and the total intensity time-averaged for 20 beads in each environment. The resulting means and standard errors are plotted in Figure 2.32. As a control to demonstrate the size-exclusion column's efficacy, the experiment was performed both for DNA/YOYO-1 complexes and YOYO-1 free in solution in the same concentration. In Figure 2.32a it can be seen that prior to the size exclusion column purification the YOYO-1 free in solution is nearly as bright as the DNA/YOYO-1 complexes. This is despite the published fluorescence boost of three orders of magnitude demonstrated by YOYO-1 upon binding to DNA. The explanation for this is that the fluorescence boost is primary due to steric factors – the individual oxalose yellows are not free to rotate upon binding to DNA and as such cannot release energy except by fluorescing. The same process may occur also upon YOYO-1 binding

to the surface of the antiDIG bead, and so the YOYO-1 may experience a similar fluorescence boost even without DNA present.

Figure 2.32b shows a dramatic drop in fluorescence under both conditions following use of the size exclusion column, reducing the brightness of the antiDIG beads with YOYO-1 but no DNA present by nearly an order of magnitude, showing that the size exclusion column was successful in removing the majority of YOYO-1 free in solution. The sample with purified DNA-YOYO-1 complexes remained approximately $30\times$ brighter than the sample without DNA, indicating that the YOYO-1 conjugated to the DNA was not removed during purification but remained on the DNA.

This was further demonstrated by performing a fluorescence "sandwich assay" following work previously done in the group [119]. Briefly, 5 $\mu$L of fluorescent sample was dropped onto a microscope slide onto which was dropped a plasma-cleaned cover slip. The falling cover slip pushes the fluid to the edges of the cover slip and this flow stretches out the DNA. Meanwhile, the low volume means the DNA is confined to a fluid layer of a few $\mu$m, allowing it to be imaged on the surface of the cover slip without needing to functionalise a cover slip and surface-immobilise the DNA itself.

In this experiment, the DNA ends are not tethered and there is no flow present. As a result, the DNA acts as an entropic spring and is expected to appear as an approximately circular fluorescent spot, as opposed to an extended line seen in dumbbell assays. The expected root mean square end-to-end distance for a wormlike chain acting as an entropic spring can be approximated as $\sqrt{< D^2_{end-to-end} >} = \sqrt{2L_C L_P}$ where $L_C$ is the contour length of the molecule and $L_P$ is its persistence length. For the DNA tether under consideration, $L_C \approx 4.8\,\mu\text{m}$ and $L_P \approx 50\,\text{nm}$ giving $\sqrt{D^2_{end-to-end}} \approx 700\,\text{nm}$. The imaged spots should therefore be approximately this size, although may be larger due to diffraction.

The sample was imaged using 1.3 mW laser power, 40 ms exposure, and EM gain set to an absolute value of 300, in a buffer without free radical scavenging. The emCCD was cooled to -80°C to reduce the dark current, which is a source of noise where electrons are excited thermally into the conduction band, and are subsequently multiplied as if the initial excitation was due to light. Representative images from this experiment can be seen in Figure 2.33. The DNA is still clearly visible above the fluorescence background, confirming that the YOYO-1 is still present in sufficient quantity even after the size exclusion column purification. The spot width is seen to be approximately 700 nm in line with expectations from the wormlike chain model,, indicating that it is indeed single DNA molecules that are imaged and the size exclusion column purification has not removed the intercalated dye molecules from the DNA construct.

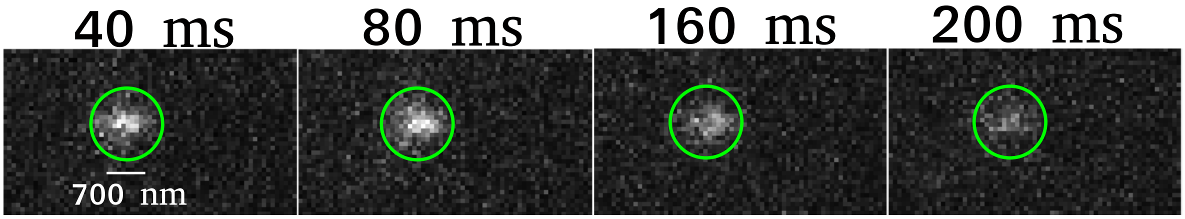


Figure 2.33: Fluorescent images of DNA/YOYO-1 purified with a size exclusion column. Because the DNA is not tethered at each end, it acts as an entropic spring and the DNA appears as a roughly circular spot instead of an extended line. An entropic spring of contour length $L_C$ and persistence length $L_P$ has expected root mean square end-to-end distance of $\sqrt{2L_PL_C}$, which for this DNA construct ($L_C \approx 4.5\,\mu$m, $L_P \approx 50\,$nm) is approximately 700 nm, consistent with the size of the fluorescent spot seen. The DNA photobleaches over time, which is consistent with expected DNA/dye complex behaviour. The DNA position is indicated by the green circle. Images were taken with absolute EM gain of 300, camera cooled to -80°C, exposure time 40 ms, total excitation laser power 1.3 mW, in buffer conditions of 50 mM Tris-HCl pH 7.6 with 10 mM $MgCl_2$.

## 2.10 Fitting the Wormlike Chain

The force-extension curves for dsDNA have been extremely well characterised, as discussed in Chapter 1. The wormlike chain model of stretching DNA at low forces is given by [74]

$$F = \frac{k_BT}{P}\left\{\frac{1}{4}\left(\frac{x}{L_C}\right)^{-2} - \frac{1}{4} + \frac{x}{L_C} - 0.8\left(\frac{x}{L_c}\right)^{2.15}\right\} \tag{2.31}$$

where $x$ is extension, $L_C$ is contour length, $k_B$ is the Boltzmann constant, $T$ is temperature, and $P$ is persistence length, a measure of the stiffness of the DNA molecule. Typical values are 40-60 nm.

In this work, the fitted parameters are contour length and persistence length, leaving $k_B$ and $T$ as constants. However, this assumption can be faulty in the case of temperature. The trapping laser attenuated to 80 mW is focussed at the sample to a radius of 1 $\mu$m, and despite water's low absorption around 1064 nm there is a local heating effect. This is further compounded by the variation in magnetic content in the 3 $\mu$m magnetic beads. The local heating due to a trapped non-absorbing silica or polystyrene bead has been measured by others to be approximately 1 K per 100 mW of trapping light [142], which was largely attributed to heating of the solution around the bead as opposed to absorbance in the bead itself. At the other extreme, gold nanoparticles were measured to heat up to 26 K per 100 mW [137]. The distribution of absorptive magnetic materials is also likely to play a contributing role – in this work, it was found that magnetic beads with a central magnetic core were not trappable while those with a ferrite shell around a styrene copolymer were. Estimating the temperature due to the trap is therefore made very difficult – as shown in Table 2.1 the beads used are highly heterogeneous in terms of ferrite content. For the purposes of this work, the temperature was estimated to be 300 K in the vicinity of the bead. As was discussed in Chapter 2, the QPD responses were normalised by dividing by the sum of the QPD signal and subtracting the average signal, which was obtained collecting QPD data from the bead trapped in solution with no DNA tether. This normalisation data was also used to find the force constant of the trap for that particular bead.

## 2.11 Experiment Protocol

The heterogeneity of the beads also cast doubt on the possibility of using a population average of trapping stiffnesses to use when plotting force-extension curves. To investigate this, 83 beads were trapped at the correct height to form a horizontal DNA tether – approximately 2.5 $\mu$m – which was calibrated by trapping the magnetic bead and bringing a surface-immobilised antiDIG bead into focus in the same field of view.

As is seen in Figure 2.34, the distribution of trap stiffnesses is spread over a considerable range. The distribution of magnetic content and bead radii should be Gaussian processes because they are random fluctuations around a mean value, and so using scipy's `scipy.optimize.curve_fit` routine, a Gaussian profile was fitted to the probability density of trap stiffness. With a mean of 4.7 pN$\cdot\mu$m$^{-1}$ and a standard deviation of 1.6 pN$\cdot\mu$m$^{-1}$, the variation of beads precludes the generation and use of reliable average trap stiffness values.

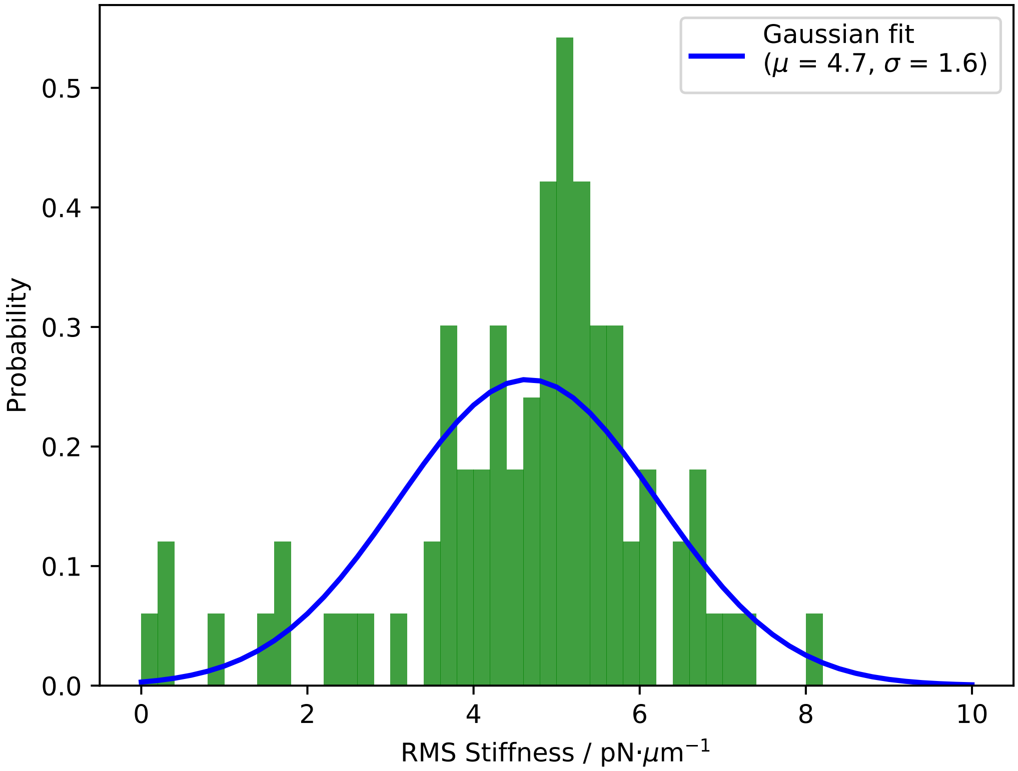


Figure 2.34: Stiffness of trap for 83 magnetic beads approximately 2.5 $\mu$m from the cover slip surface, fitted to a Gaussian distribution. It can be seen that the standard deviation of trap stiffness is high compared to the mean and so stiffness should be calculated for each individual bead used.

Therefore, a protocol was developed to ensure that each tether was a single DNA molecule, and to find that bead's unique force constant, before beginning a full experiment. Briefly summarised, a full experiment was performed as follows:

1. Trap a freely-floating magnetic bead
2. Bring it to the correct height above the cover slip, using a surface-immobilised antiDIG bead as a reference
3. Collect QPD data to use for normalisation and to find trap stiffness
4. Generate a tether by bringing the beads within 500 nm of each other using the nanostage and waiting 2 minutes
5. Oscillate the nanostage and collect QPD data to generate a force-extension curve

6. Stop the oscillation and apply a known force using the force clamp
7. Under- or overtwist the DNA at 2 Hz using the Helmholtz coils, counting the number of rotations applied
8. Stop the rotation, hold $\sigma$ fixed through application of a DC voltage across the Helmholtz coils, and allow the force clamp to equilibrate
9. Reverse the rotation, applying the same number of turns as were applied previously
10. Repeat step 8

Data analysis was carried out after the fact using bespoke software written in Python 2.7 using SciPy, NumPy, and Matplotlib.

## 2.12 Summary

In this Chapter, the principles, design, construction, verification, and method development for the magneto-optical tweezers and fluorescence microscope were presented, and the steps for the final experimental protocol were explained. The imaging capabilities of the microscope were described, and the considerations and compromises which must be made in choosing fluorescent dyes have also been discussed and the choice of YOYO-1 justified in this context.

The mathematical theory behind optical tweezing has been given, and discussed in the context of real-world situations and difficulties. The original work done to characterise the optical trap and optimise experimental conditions has also been presented. The magnetic beads were trapped and their stiffnesses quantified. The QPD response and trap stiffness with height were found to be in good agreement with theoretical predictions, giving confidence to the collected QPD data. Analysis of the stiffness probability distribution showed that a Gaussian with mean 4.7 $\mathrm{pN \cdot nm^{-1}}$ and standard deviation 1.6 $\mathrm{pN \cdot nm^{-1}}$ fit the data, demonstrating that the variability between beads is such that each individual must be calibrated – in other words, a population average stiffness would be too inaccurate and incorrect forces would be applied.

The principles behind magnetic tweezing with Helmholtz coils have been described, and the proof-of-concept experimental data presented to determine that the machine is capable of applying torque and force to magnetic beads.

The work done to create an experimental protocol was given. It was found that a high ionic strength buffer enabled reliable adherence of the 4-4.9 $\mu$m diameter antiDIG bead to a nitrocellulose-treated surface. The methods for creating the DNA constructs were detailed, and agarose gel electrophoresis was used to verify that the PCR has been successful. Agarose gel electrophoresis of the constructs in the presence of molecular antibodies demonstrated that the functionalised nucleotides had been successfully incorporated into the constructs' handles.

The incubation time for the 4 kbp tether was found by way of a fluorescence assay monitoring total antiDIG bead intensity through time, and a significant jump in fluorescence after 30 minutes was used as indication that DNA conjugation had been successful. The 15 kbp tether was imaged directly with epifluorescence microscopy to verify that it has formed a complex with the immobilised antiDIG bead. Tethers were formed by trapping a free magnetic bead and holding it within a few hundred nm of the antiDIG bead for a period of 2 minutes. Tether formation success was triaged using the nanostage to observe qualitatively whether the bead was being pulled from the trap at extensions near to the expected contour length. If a single tether was indicated, the DNA molecule was used for production experiments which will be presented in the next Chapter.

Epifluorescence microscopy conditions were investigated. Free YOYO-1 in solution was found to bind non-specifically and increase intensity of the beads such that the DNA was not visible. A column purification was used to remove excess dye and its performance found to be excellent. All fluorescently labeled DNA in this thesis has undergone this column purification. The effect on the dyes of the imaging buffer were quantified and it was found that the presence of 2 mM Trolox was deletrious in these conditions. Therefore the imaging buffer used for later experiments consisted of the oxygen scavenging system without the triplet state quencher. In order to keep results as comparable as possible, this imaging buffer was also used in experiments with no YOYO-1 present.

The functionality of the force clamp was demonstrated by applying a target force to a tethered optically trapped bead. The resulting position trace shows the smooth movement of the force clamp towards the target, and a controlled approach in order not to overshoot it.

Finally, the methods to fit the wormlike chain were given, and the work of this Chapter was combined to give the final experimental protocol which was used to generate the data in the following Chapter.

# Chapter 3

# DNA Manipulation *in vitro*

In this Chapter, the results of applying the magneto-optical tweezers to double-stranded DNA (dsDNA) will be presented, as will the fluorescence images obtained of DNA tethers. The experiments and analysis in this Chapter were performed by the author.

## 3.1 Brightfield Assays

### 3.1.1 Force-Extension Behaviour

The force-extension curves for unlabelled dsDNA were generated according to the protocol in Chapter 2. The nanostage was oscillated in a triangular wave as in Figure 3.1a). The position of the nanostage is not, however, the same as the extension of the DNA itself. The DNA will pull the bead from the centre of the trap at high extensions, so that the extension of the DNA is less than would be obtained if looking only at the nanostage.

To correct for this, the voltage to position conversion as determined in Chapter 2 is used. The position from the centre of the trap is seen to be

$$\Delta = k_S \frac{V_y}{V_{sum}}$$

where $k_S$ is the trap sensitivity in $\mu\text{m} \cdot V^{-1}$, $V_y$ is the QPD response along the vertical axis and $V_{sum}$ is the total QPD response. It trivially follows therefore that the total DNA extension is given by

$$L_{DNA} = y_{Nanostage} - y_0 - \Delta \tag{3.1}$$

where $y_{Nanostage}$ is the position of the nanostage and $y_0$ is its position at $L_{DNA} = 0$.

Then the force is clearly

$$F = -k_y L_{DNA} \tag{3.2}$$

where $k_y$ is the force constant along the stretching axis.

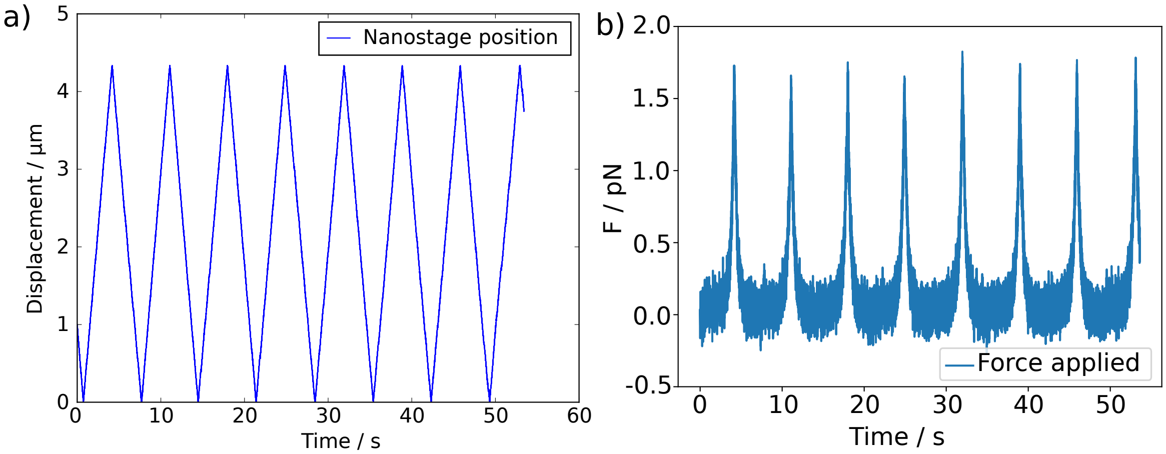


Figure 3.1: a) Triangular movement of the nanostage for generating force-extension curves. Distances are relative to the two beads touching and the DNA extension therefore being 0. The upper limit of the extension was determined experimentally: the oscillation beginning with low amplitude and growing until the bead was being visibly pulled from its equilibrium position. b) Force applied to a DNA tether throughout a force-extension assay. The sharp peaks show the extreme of force applied to the DNA, while the smaller peaks around the equilibrium position are caused by interference in the trap from the surface-immobilised antiDIG bead which forms the anchor for the DNA tether

The result of applying this is seen in Figure 3.1b). Rather than smooth peaks, the force is low for much of the extension and rapidly increases as the increased entropic force on the DNA pulls the bead out of the trap. The horizontal axis in Figure 3.1b) is time, thus showing each individual oscillation. In order to fit the wormlike chain, the data is overlaid by plotting the force response of the DNA against total DNA extension, which has a similar effect to averaging the cycles: if a point is noisy the effective weight of those points in the curve fit decreases.

Plotting the nanostage extension against the force gives rise to force-extension curves as shown in Figure 3.2. The fitted values of persistence length, between 35 and 70 nm, are in the correct range, though have a wide spread. The contour lengths found consistently exceed the 4.5-4.8 $\mu$m expectation value for the DNA midsection.

There is however one more correction that should be applied. The data gathered above is the force along the axis of nanostage oscillation. However, it is not guaranteed that the tether lies perfectly along that axis – in fact, it is intuitively more likely that the tether would lie off-axis and therefore contribute to the total force. Two methods were tried to correct for this, which will be referred to as “force inference” and “absolute force”.

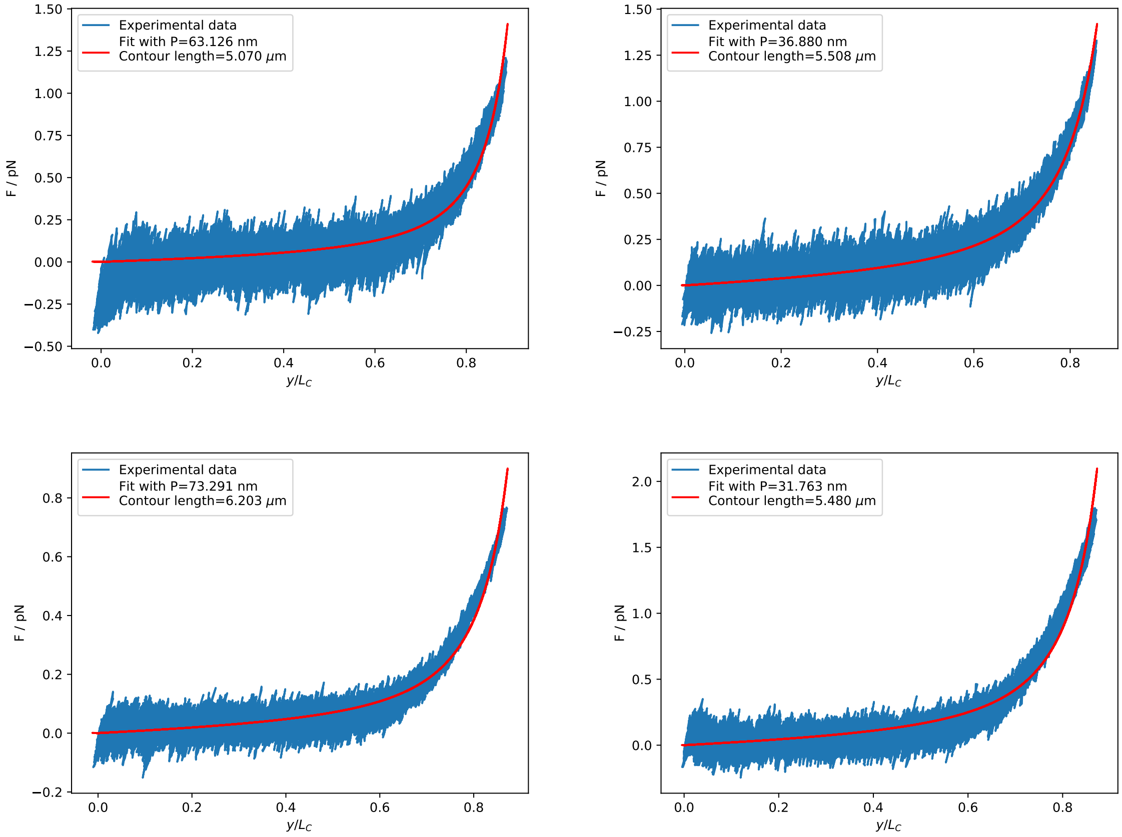


Figure 3.2: Four wormlike chain fits to uncorrected force-extension curves. The fitted values for persistence length are of the correct scale, but the contour length is consistently longer than would be expected

In order to perform the force inference correction, the total force is inferred from the angle $q$ between the DNA tether and the axis of oscillation by noting that the force triangle consisting of total force, $F_x$, and $F_y$, and the distance triangle given by the total displacement, $d_x$ and $d_y$ are similar. The angle $q$ between the nanostage oscillation axis and the tether is then

$$q = \arctan\left(\frac{F_x}{F_y}\right) \tag{3.3}$$

where $F_x$ is the force in the axis perpendicular to the oscillations and $F_y$ is the force along the axis of oscillation. Then from trigonometry we can correct the displacement and force measured:

$$d = \frac{d_y}{\cos q} \tag{3.4}$$

$$F = \frac{F_y}{\cos q} \tag{3.5}$$

where $d_y$ is the nanostage displacement.

Since this data (Figure 3.3) is extremely noisy, the raw value of the angle $q$ cannot be used. Instead a polynomial fit was applied to smoothly approximate the behaviour of the $q$ angle.

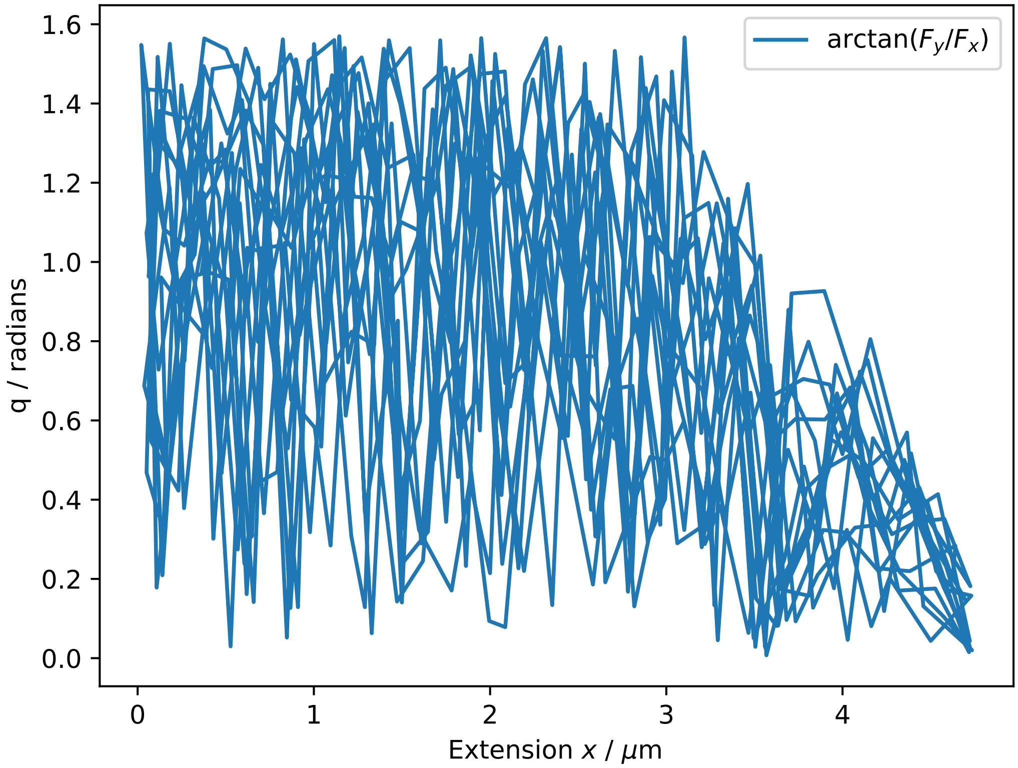


Figure 3.3: Example $q$ angle as a function of displacement calculated as $\arctan(|\frac{F_x}{F_y}|)$

One force-extension plot corrected through force inference and fitted with the worm-like chain is seen in Figure 3.4. The shape of the force-extension trace has distorted and no longer resembles a smooth stretching curve that would be expected. The persistence length is very close to previously found values, but the contour length remains longer than would be expected for 14 kbp B-DNA.

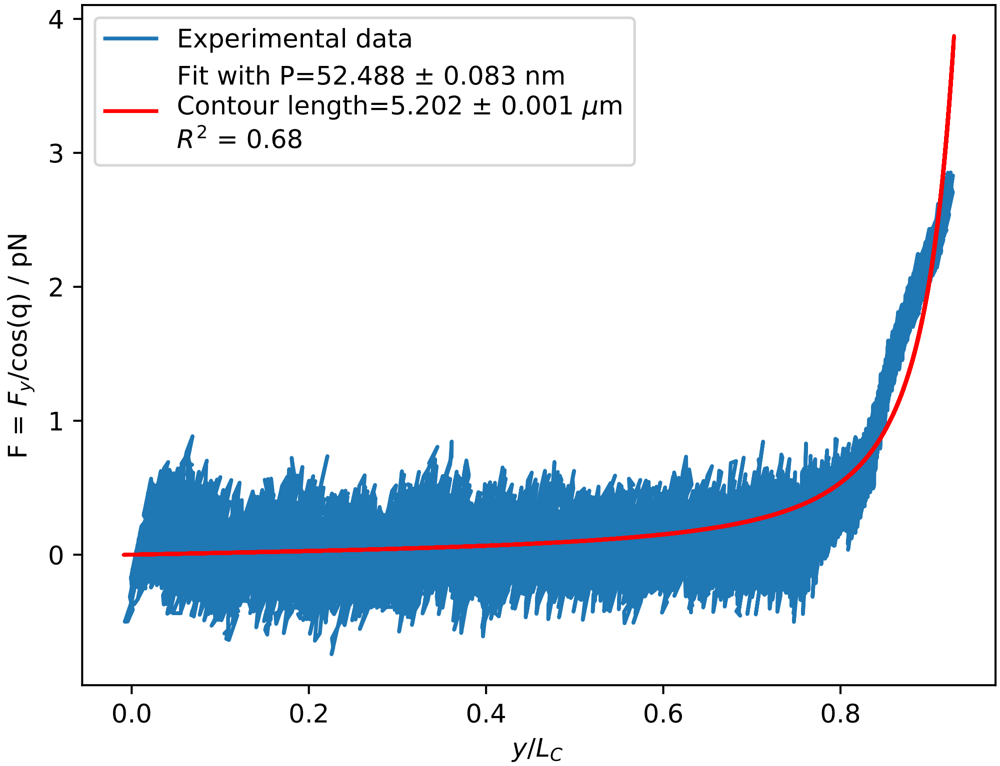


Figure 3.4: A force-extension curve corrected through force inference. The persistence length remains at a value that would be expected, while the contour length remains above the theoretical length of 14 kbp B-DNA, 4.2 $\mu$m.

The second method to correct for an off-axis tether is the absolute force method. The $x$ and $y$ components of force are known so the total force is trivially found with the Pythagorean theorem. The noise on this measurement will be lowest at large extensions as more force is applied directly on the bead. The angle $q$ is most reliably found through Equation 3.3 at high extensions. The absolute displacement in the $x$ direction is then found at large $d_y$ as

$$d_x = d_y \tan q \tag{3.6}$$

thus

$$d = \sqrt{d_x^2 + d_y^2} \tag{3.7}$$

for the constant $d_x$ and all nanostage positions $d_y$. The total force is simply

$$F = \sqrt{F_x^2 + F_y^2} \tag{3.8}$$

Plotting the absolute value of force means that the fluctuations are never negative for low displacements, and so the mean of the force cannot be 0. This can be heuristically corrected by subtracting the mean force at low extension. The result of performing this procedure is given in Figure 3.5. The mean at low extension is now 0 as would be expected for the system. However, the fitted contour length has not reduced from previously found values.

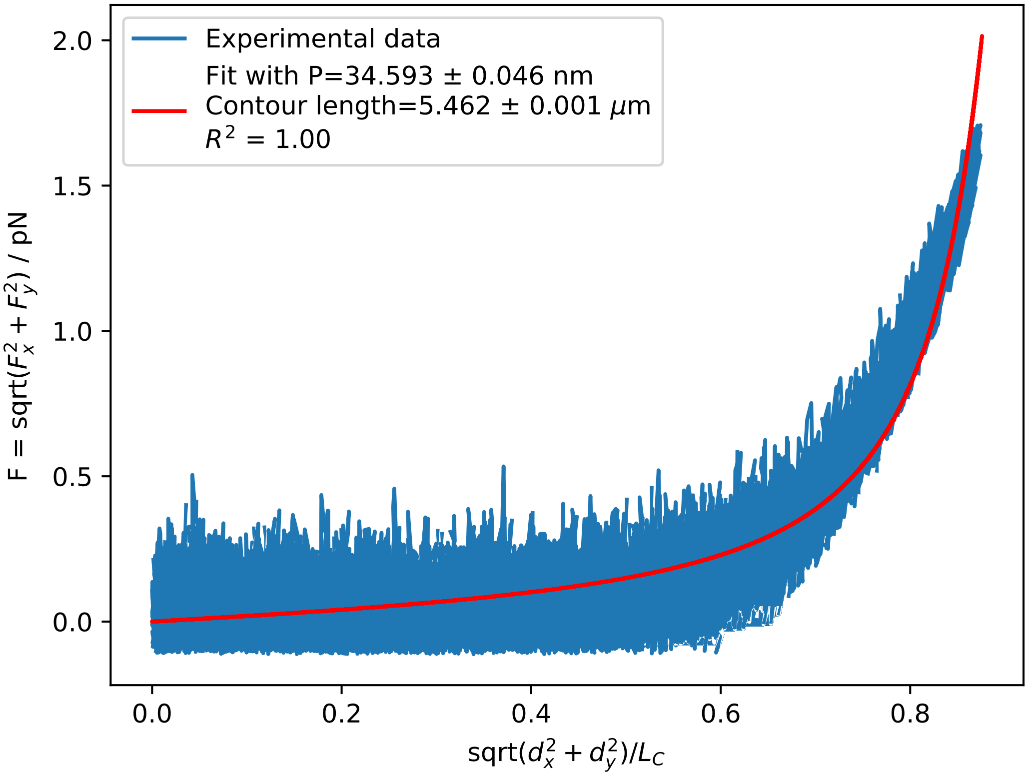


Figure 3.5: A force-extension curve corrected with the absolute force method with the mean force at low extension subtracted. Though the mean at low extension is now 0, the fit parameters are still not in line with expectations.

Across the methods, the persistence length has consistently been found to be in the correct range of values, though the spread of data was high. More repeats are needed to reliably find the persistence length under these experimental conditions.

By contrast, the fitted contour lengths have consistently been in disagreement with expectations, and have been found to be in the range $5 < L_C < 7\,\mu$m. Rescaling the force and displacement data using the methods presented did not give improved estimates of the contour length. Further, rescaling forces and displacements through division by $\cos q$ led to the graph profiles being distorted and the fits themselves being unstable and failing to converge. Correcting via the absolute force method resulted in a more conventional looking plot, but the $L_C$ values found did not improve. Although it appears not to have had an effect, this correction was in fact deletrious – as the mean force was no longer 0 at low displacements, the forces needed to be heuristically corrected by subtraction, which impacts the force seen at high displacement. Since both corrections directly led to poorer fit performance and incorrect force-extension profiles, the force extension graphs for DNA/YOYO-1 complexes were not rescaled.

The fit values, with $30\,\mathrm{nm} < L_P < 75\,\mathrm{nm}$ and $5 < L_C < 7\,\mu\mathrm{m}$, are seen consistently across the experiments dealing with DNA both with and without YOYO-1. The persistence lengths are in agreement with previously found results in literature, while the 4.2-4.8 $\mu$m predicted contour length has not been found. There are multiple reasons why the contour length may not be found to be as expected:

**Poor quality data** A faulty QPD could easily produce data which was noisy or unusable. However, the QPD itself was regularly checked, and its response function plotted as in Figure 4.8. Good agreement was consistently seen between the QPD's actual and theoretical curves, and the fitted persistence length was in line with expected values.

**Software malfunction** Fitting the wrong functional form, or manipulating the DNA wrongly during fitting could lead to subtle and difficult-to-detect bugs including an overestimation of the contour length. However the distances calculated with the code are consistent with those seen in the brightfield images.

**Nicked DNA** DNA with breaks in the covalent bonds along their backbones behave very differently from torsionally constrained DNA [93] which could account for the unusual results, as the DNA could become in effect two stretches of dsDNA with a ssDNA linker region. But if a DNA molecule were torsionally unconstrained, it would not ever reduce in length due to applied rotation, and the fitted persistence length would not be close to expected values. Furthermore, T4 DNA Ligase was used in all imaging buffers to repair nicks and guard against this possibility. The routine observation of shortening due to rotation suggests that the DNA itself was intact.

**Those are the correct values** Most possible errors in either the software or the DNA itself are easily accounted for in the preliminary method development work seen in Chapter 2, and the results themselves, while unusual, are relatively consistent.

> It is known that experimental conditions affect mechanical properties of DNA – for example, salt concentration affects DNA persistence length, as does salt type (monovalent or divalent) [70, 143]. However, the picture in these experiments is more complicated than the published work: the buffer contains divalent ions, ATP, DTT, ligase, catalase, glucose oxidase, glucose, and BSA. As a result, the screening behaviour and effect on the DNA mechanical properties is largely unknown. It is possible that the buffer used led to the DNA adopting a more extended conformation than would usually be seen.

### 3.1.2 Force Clamp Assays

First, the unlabelled DNA molecules were subjected to the twisting assay as described. Previous studies using magnetic tweezers based on permanent magnets have shown that there exists a transition in the behaviour of undertwisted DNA. Strick *et. al.* [92] demonstrated that when DNA is held with a tension above a threshold in the range 0.6-0.8 pN, DNA undertwisted to values in the range $-0.05 \leq \sigma < 0$ did not reduce in contour length, while when overtwisted the DNA did shorten.

The response to applied $\sigma$ under a tension of 0.55 pN found in this work is seen in Figure 3.6. The DNA contour length decreases slowly and linearly as sigma reduces from 0 to around -0.015, and there the rate of shortening increases but stays linear until the rotation ceases, qualitatively in agreement with previous studies. Over the subsequent equilibration period (indicated by an arrow and a dashed line) the length of the DNA decreases yet further until it is apparently negative. However, this is an artifact from the two beads being so close together that the surface immobilised antiDIG bead interferes with the trapping laser.

Upon retwisting, the DNA follows a different, more shallow path. The extension increases slowly at first but again accelerates and is linear until $\sigma = -0.09$ where there is another transition to a shallower linear response. Together, these paths appear to constitute a form of hysteresis curve, in which the DNA is following different paths dependent on its starting configuration.

This hysteresis-like behaviour continues as the tether is overtwisted. Again, the reduction in extension is slow at first but accelerates, with the overall extension behaviour thus appearing to be two linear regions. The equilibration stage reduces the extension yet further, and the relaxation stage once again follows a path distinct from the one taken during the overtwisting.

The response of the force clamp was then assessed for this procedure by plotting the force against time. Ideally, the resulting graph would be perfectly flat. However, it can be seen in Figure 3.7 that there are peaks and troughs in the different undertwist, overtwisting, and relaxation stages. The peaks in the twisting stages correlate with the changes in gradient in Figure 3.6, demonstrating that a spike in force produced by the DNA directly causes a more rapid change in stage position, and that the force clamp is working as it should.

Further evidence of the force clamp working as it should is also present in Figure 3.7. The deviation from the target force (0.55 pN) is low at around 0.3 pN, which is a good level of accuracy compared with other force clamps. The overall drift over the full experiment is also extremely small. A straight line fit to the force data shows that the drift is 0 pN per second to three decimal places.

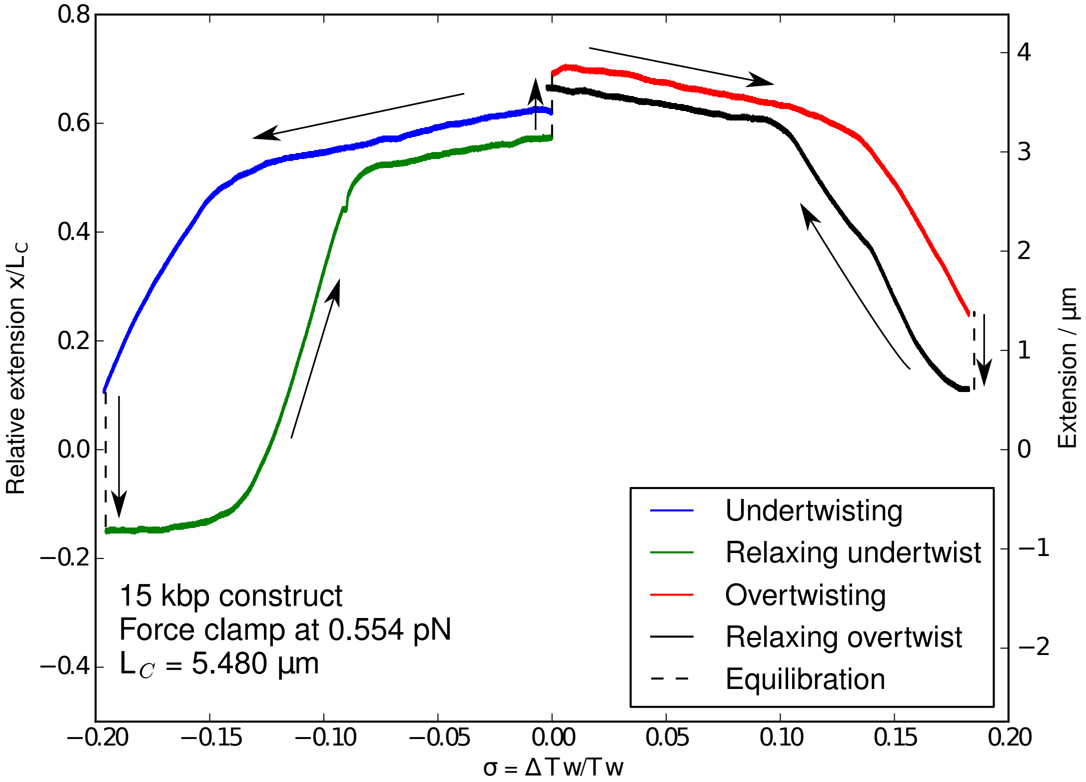
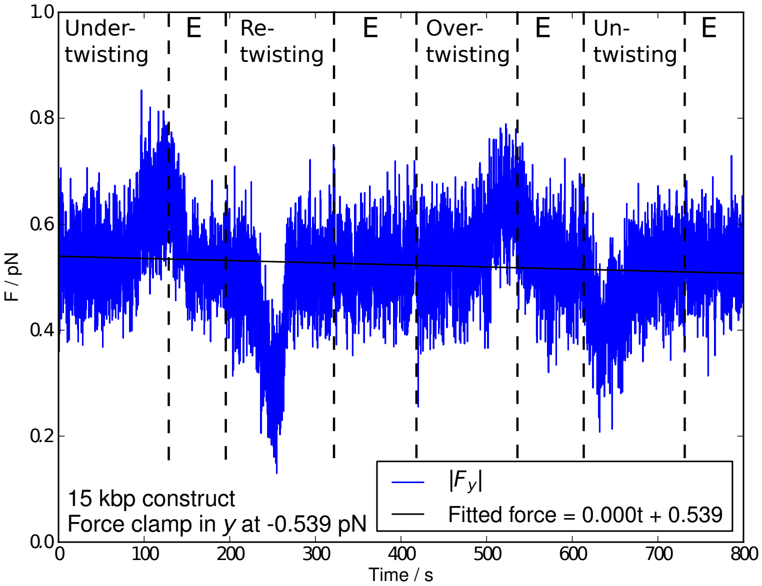
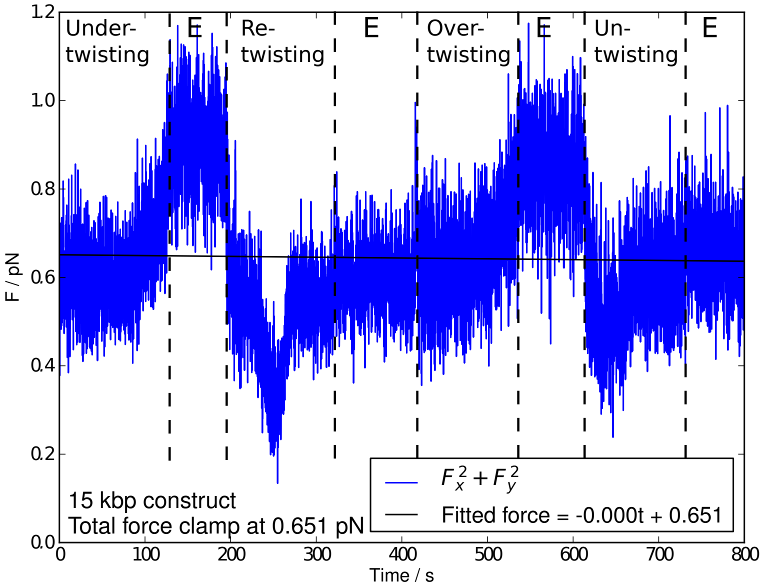


Figure 3.6: Twist-extension curves for a single DNA tether under low (0.55 pN) force. Although the applied force is just above the threshold seen in [92], significant shortening of the DNA is observed with undertwisting. The overtwisted DNA is seen to shorten as expected. In both cases, the under and overtwisted DNA continues to shorten in length after twisting has ceased; this equilibration is indicated by dotted lines. When being returned back to its canonical twist, the DNA follows a different path reminiscent of a hysteresis loop. The direction of travel around this loop is indicated by arrows.



a)



b)

Figure 3.7: a) Force through time for a single DNA tether undertwisted and overtwisted with the force clamp set to 0.55 pN in $y$. The force applied remains within 0.3 pN of the target force for the majority of the trajectory. b) The total force applied to the bead is shown for the twisting assay. Due to an off-axis tether, the total force is higher than the force targeted by the force clamp, which works only in 1 dimension. However, the difference in forces is only 0.1 pN, and when plotting this total force it is kept within around 0.5 pN of the initial force, indicating that the off-axis tether acts as a constant force offset rather than a dynamic source of error. For both a) and b), overtwisting, undertwisting, and relaxation regions are marked and E indicates an equilibration stage, *i.e.* no rotation is applied to the magnetic bead and it is held torsionally fixed with a DC voltage across the Helmholtz coils.

However, as there may be an off-axis component of force, the applied force clamp is not necessarily the correct measure. Shown also in Figure 3.7 is the absolute force experienced by the bead. The total force that is clamped is now around 0.65 pN – 0.1 pN higher than purely in the $y$ axis, but the force once again stays within 0.5 pN for the majority of the experiment. This combined with the smooth response seen in Figure 2.27 show that the force clamp implementation is correct for the $y$ dimension, but it is important to make sure the tether lies properly along the $y$ axis. This would minimise data post-processing, remove the need for using both $x$ and $y$ stiffness and sensitivity, and hence maximise the accuracy of the measurements.

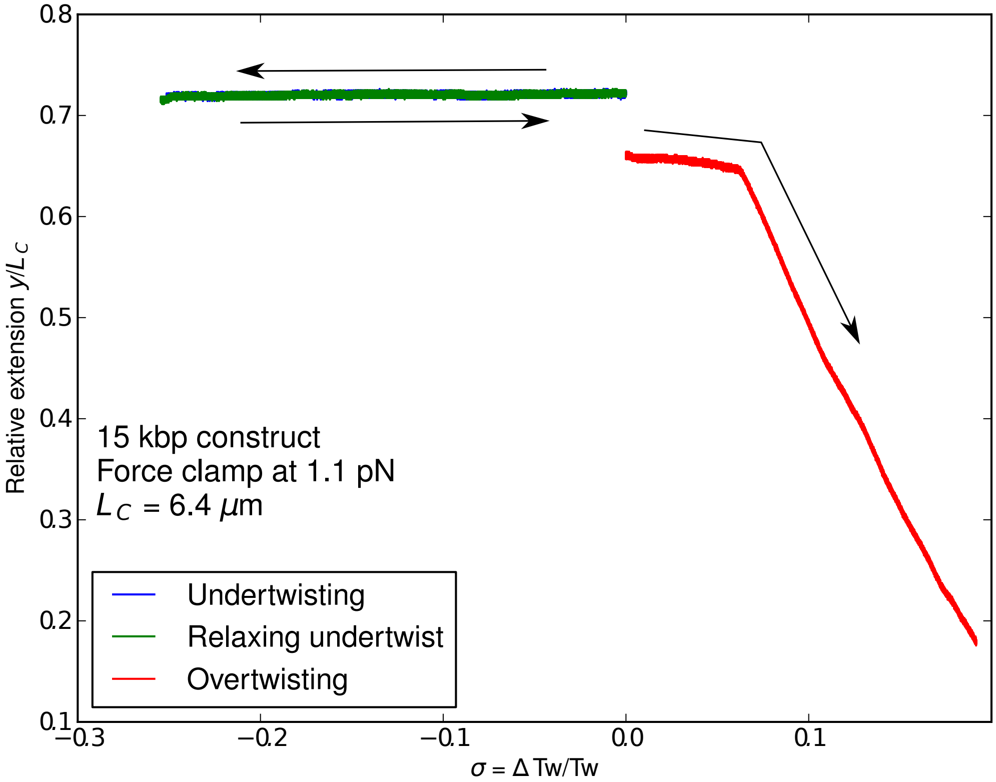


Figure 3.8: The twist-extension behaviour for a single dsDNA molecule under 1.1 pN of tension. When undertwisted, the DNA does not shorten appreciably even for extreme values of $\sigma$, and returns to its canonical twist structure along the path it took previously. Overtwisted DNA is seen to initially plateau, with no shortening seen for low values of $\sigma$, but then shortens dramatically. This behaviour is in agreement with previous work [92]

Finally, the force clamping assay was performed for DNA held fixed with a higher force – this time, 1.1 pN. The results, seen in Figure 3.8, are strikingly different from those in Figure 3.6. The undertwisting section, the direction of movement again indicated by arrows, is as expected from literature [92] as it is above the force required to prevent shortening of a tether due to undertwisting at this $\sigma$. The DNA does not reduce in length at all, and returns by exactly the same path to its original configuration as expected.

When overtwisted, the DNA plateaus and does not shorten until the twist reaches a critical value whereupon shortening happens quickly and monotonically. In this experiment, the part of the curve which dealt with the untwisting of the twisted DNA construct was not usable, as part way through the untwisting the DNA or DNA-bead conjugation broke, and as a result the full cycle as in Figure 3.6 could not be completed. However, the graph shows the results in good agreement with the $\sigma - x$ curve in [92]. Sadly the presence or otherwise of hysteresis could not be evaluated in the overtwisted region, however the undertwisted region appears to follow the same path in both directions.

Both the higher and lower force regimes clearly show corners in the twist-extension curves. These demonstrate that there are distinct regimes in the conformational changes the DNA undergoes in both the forward and reverse twisting direction [5]. The cause of this is the change from a twist dominated system to a writhe dominated one [144]. At the buckling point, the introduced twist is absorbed by formed plectonemes instead of being absorbed by the DNA double helix overtwisting [144]. In this work, that process is seen in reverse as well, and the differences in twist-extension traces followed suggest that the DNA follows different structural pathways in each direction – a form of structural hysteresis.

The data in Figures 3.6 and 3.8, at an applied force of 0.55 pN and 1.1 pN respectively, showed good agreement with literature. Previous work showed that the critical force for preventing undertwist-induced shrinking is between 0.6 and 0.8 pN [92]. The magneto-optical tweezers have therefore found the two different force regimes previously identified, with the exact limiting force in the expected range. However, the length reduction for $F = 0.55$ pN was more dramatic than seen previously, likely due to buffer conditions. The response upon overtwisting was in good agreement with previously published research [92]. Further experiments will need to be undertaken to determine the exact critical force for the system used, and in less complex buffers for more quantitative comparison to previous work.

Hysteresis-like behaviour can very clearly be seen in Figure 3.6. Many previous works have endeavoured to work in situations of thermodynamic reversibility and have therefore not noted hysteresis loops, but two contradictory reports have been published.

Using magnetic tweezers, researchers in the Dekker group showed plots of DNA undertwisting and retwisting without hysteresis as part of a larger study on the phase diagram of DNA structure during undertwisting [145], whereas other work by Sheinin *et. al.* [99] showed hysteresis at low forces, in common with this work. However, the hysteresis seen by Sheinin *et. al.* is of the opposite nature to that seen in this work. Their untwisting curves sit below their retwisting curves, in direct contrast to Figure 3.6. Once again, the experimental conditions are so complex that it is difficult to identify what could be the root cause. Additionally, the hysteresis behaviour could be force-dependent.

However, reasonable hypotheses can be made. The choice of buffer is again a possibility, and the first things that should be done to further this work is to perform control experiments. The second primary factor is likely to be the rotation rate. In this work 2 Hz was used, with no stopping or equilibration until the full rotation had been applied. It is possible that this rate produces novel structures when twisting, and that the perturbation needs more time to be distributed throughout the system. Further work would be needed to repeat the experiments at, for example, 0.1 Hz, 0.5 Hz, and 1 Hz in order to assess the dependence of this on the traces obtained. It is remarkable that a rotation of 2 Hz is a high enough rate to cause such effects, however, and if this measurement is repeated and verified it would have implications on DNA structure and packing, implying as it does that some effects of molecular machines and proteins on DNA may be local on timescales well beyond the typical milliseconds of cellular processes.

## 3.2 DNA/YOYO-1 Complexes

The following measurements were performed with a saturating amount of YOYO-1 to ensure that the DNA was fully labelled. Specifically, the ratio of YOYO-1 to DNA base pairs was 2:1. Although as shown before this is too high a concentration to use for fluorescence imaging, the high ratio is appropriate for a brightfield assay in which it is essential that the DNA is maximally occupied with YOYO-1 molecules.

### 3.2.1 Force-Extension Behaviour

As before, the force extension curves for the DNA/YOYO-1 complexes were found. Two examples of this can be seen in Figure 3.9. Once again, the persistence lengths are around the correct values, and the contour lengths are considerably above 5 $\mu$m The expectation for DNA with YOYO-1 present is an undertwist of the helix and an increase in contour length of around 38% but no increase in persistence length [85, 84]. While the persistence length is comparable to the DNA without YOYO-1, the contour lengths remains the same, at between 5 and 6 $\mu$m.

There are various reasons why this may be the case. Once again, an error could lie in the data gathering, the QPD, or the fitting. However, each of these has been well tested, and so it seems unlikely. Moreover, the consistency of the results suggests that any such error is systematic and does not vary much with time, while if faulty equipment or a poor analysis algorithm were the cause, it is likely that the errors would be widely scattered due to stochastic electrical malfunctions or numerical instability. The correct range found for persistence lengths further indicates that the explanation lies elsewhere. Another reason could be poor binding of YOYO-1 to DNA, but fluorescence images taken with the dye demonstrate that the conjugation is successful. Instead, the cause may well be in the imaging buffer itself. The complex mixture of enzymes, buffers, salts, and sugars is far from the tightly controlled conditions used when studying mechanical properties. In [85] the DNA and YOYO-1 were prepared in 10 mM Tris with 1 mM EDTA and deposited on a mica surface, imaged, and the mechanical properties extracted. The vast difference in buffers makes comparison with the results presented here very difficult if not impossible. Further, analysing mechanical properties by deposition can often give different results from properties seen in solution. Once again, further work using the magneto-optical tweezers in this microscope should be performed in less complex buffers to more accurately compare with other findings.

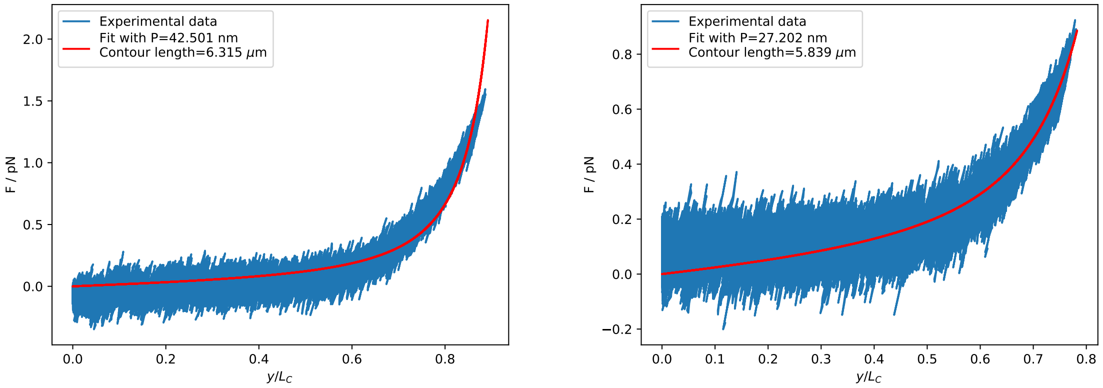


Figure 3.9: Force extension curves for DNA complexed with YOYO-1. The persistence lengths have not changed upon introduction of YOYO-1, in accordance with previous work [85]. However the previously reported 38% extension in contour length are not in evidence.

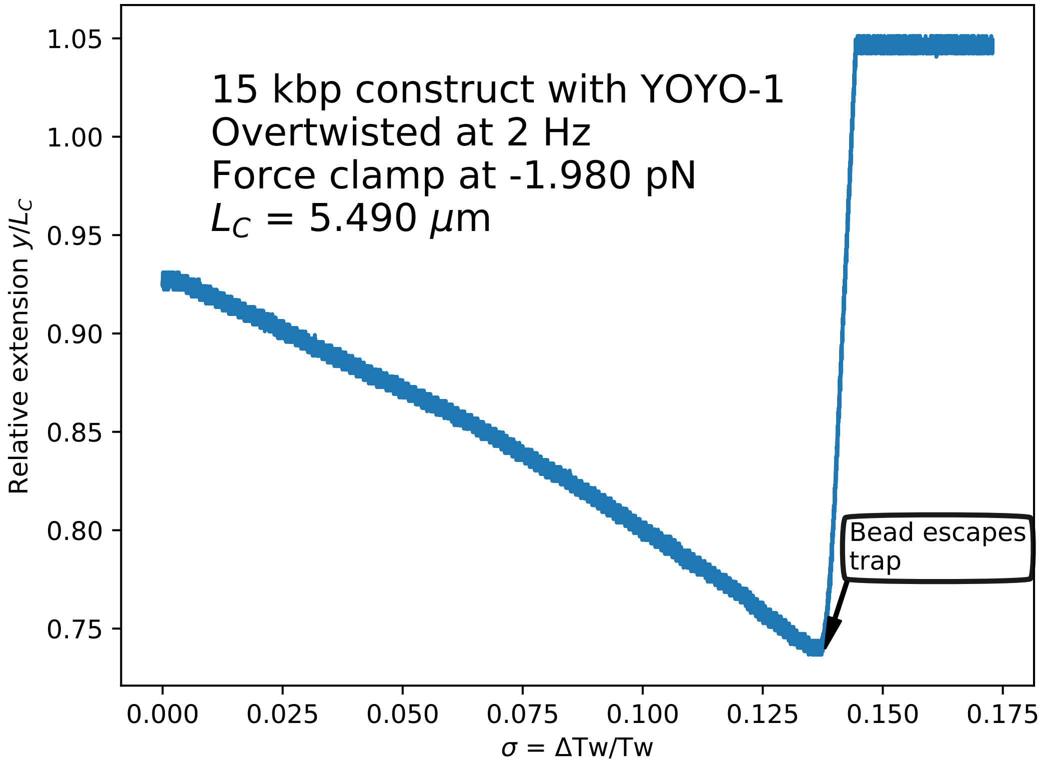


Figure 3.10: Extension with overtwisting a single DNA/YOYO-1 complex with applied 2 pN tension. It can be seen that the bead escapes the trap at high $\sigma$, suggesting the trap cannot keep the force constant.

### 3.2.2 Twist-Extension Relations

The twist-extension relations for DNA and YOYO-1 were then examined in the same way as previously. Initially, a relatively high force of 2 pN was applied to a single DNA/YOYO-1 complex. which was then overtwisted. As seen in Figure 3.10, the DNA extension drops with increasing $\sigma$ as seen previously. However, around $\sigma = 0.14$, the bead is seen to escape the trap altogether, the implication being that the force clamp does not keep up with the rate of change of force exerted by the DNA. As the force clamp has been shown to successfully maintain forces over a wide range of $\sigma$ for bare DNA, this suggests that there is a violent change in the force behaviour with YOYO-1 present.

To quantify this more, the force through time was plotted as in Figure 3.11. The force is constant initially, but at the time the bead exits the trap a sharp change in the force is seen.

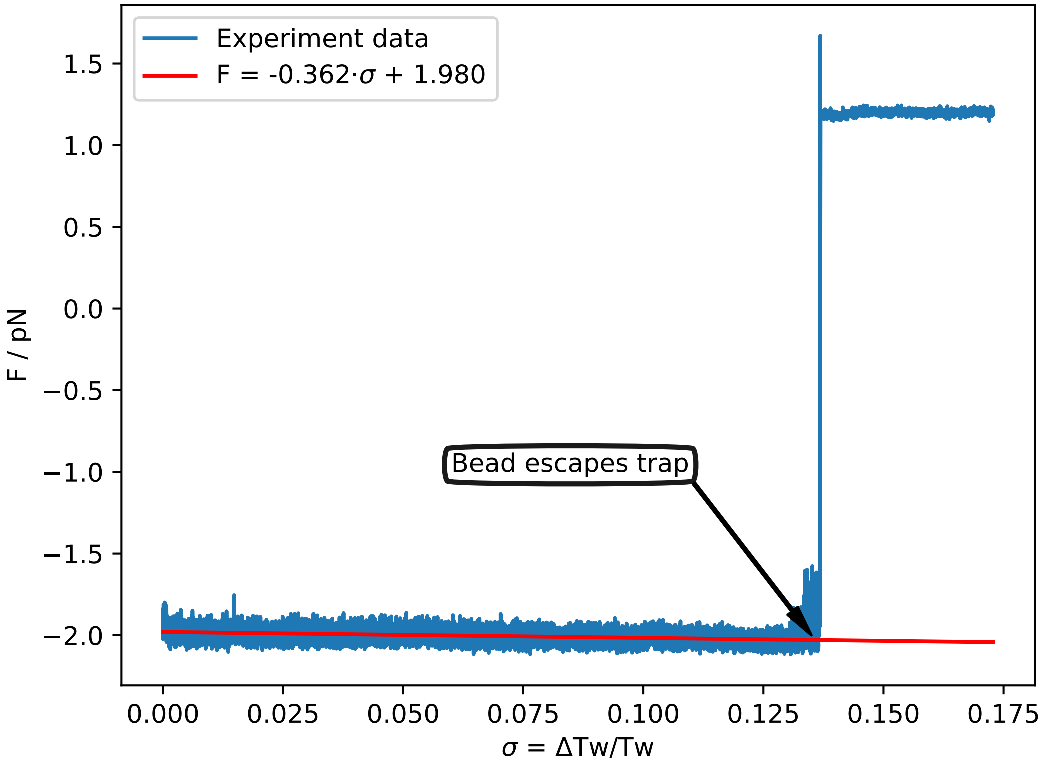


Figure 3.11: Force vs applied relative twist for one DNA/YOYO-1 complex in Figure 3.10. The force remains nearly constant over the twisting until the bead's exit, so a sharp application of force is the likely explanation.

In DNA twisting experiments, buckling behaviours are commonplace [5, 145] and it is reasonable to expect that such behaviours would be seen with YOYO-1 present also. In this work buckling points were seen in the form of gradient changes in Figures 3.6 and 3.8 and the force clamp was observed to deal with the transition well. That it is unable to do so with YOYO-1 present suggests that the transition is sharper with the intercalator present, a hypothesis which is intuitively sensible. A YOYO-1 intercalation will provide additional torsional stiffness, but beyond a certain overtwist the YOYO-1 may be ejected from the DNA. At that point, with the support gone, the plectoneme formation may begin, causing the DNA to suddenly buckle.

This hypothesised sharp transition with YOYO-1 present was tested further by performing a twist-extension experiment with a lower tension applied to the DNA, in this case 0.7 pN. The plot of extension vs applied $\sigma$ is given in Figure 3.12. Its shape is qualitatively similar to those seen for the cases without YOYO-1 in Figure 3.8. However, the gradient change occurring at $\sigma \sim 0.068$ is steeper than previously, as is a similar transition at around $\sigma$ =0.055 when removing the overtwist from the DNA. This is not due to the applied tension – as the tension applied in this case is higher than that in Figure 3.6 it would be expected from other work [94, 99, 145] that the gradient would be lower.

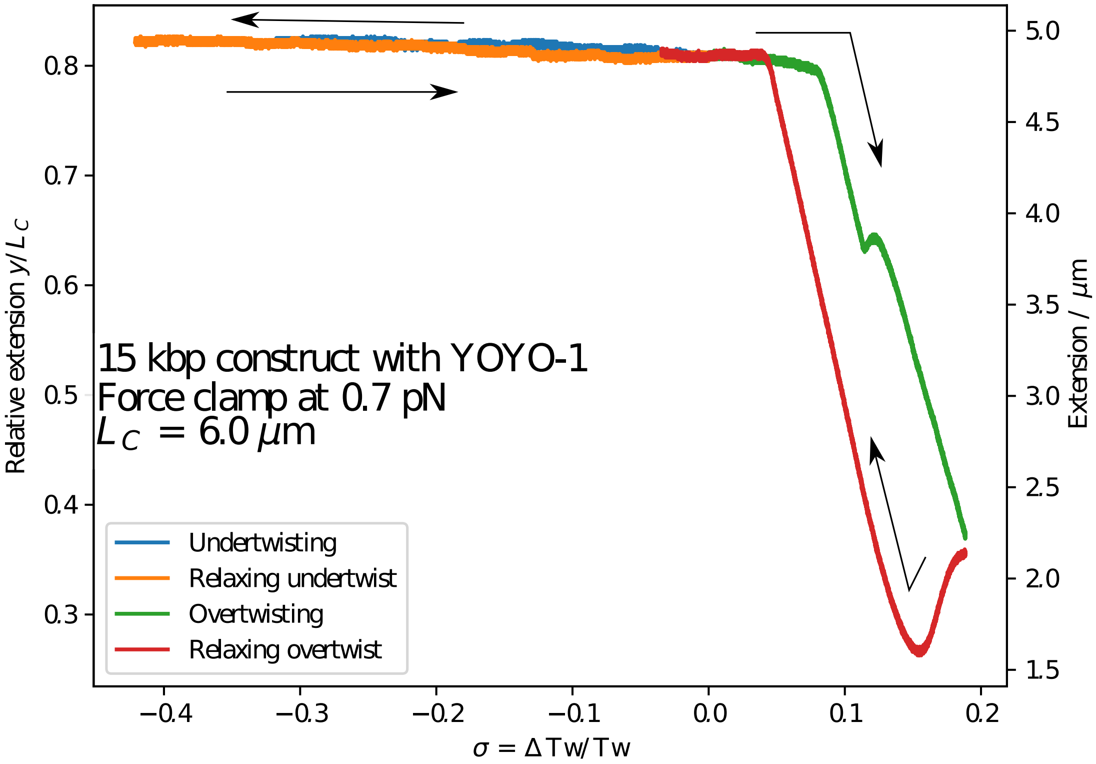


Figure 3.12: Extension vs applied relative twist for a single DNA/YOYO-1 complex with applied tension of 0.7 pN. The curves behave similarly to those in Figure 3.8. With applied undertwist, the DNA does not shorten appreciably. When overtwisted, the DNA length plateaus for some time before quickly declining. After the magnetic tweezers are stopped and the $\sigma$ therefore held constant, the DNA slightly shortens further. As it is untwisted back to its original conformation, the DNA follows a very different twist-extension path, suggesting hysteresis.

However, as can be seen in Figure 3.13, the gradient changes differ between the two cases. The two traces begin with comparable gradients of around -0.2 $\sigma^{-1}$ , but when the buckling points are reached the behaviour varies. In the case with YOYO-1 present, the DNA experiences a buckling event that changes the gradient of the $\sigma - x$ plot by 5 $\sigma^{-1}$ , 1.25× higher than that for the bare DNA.

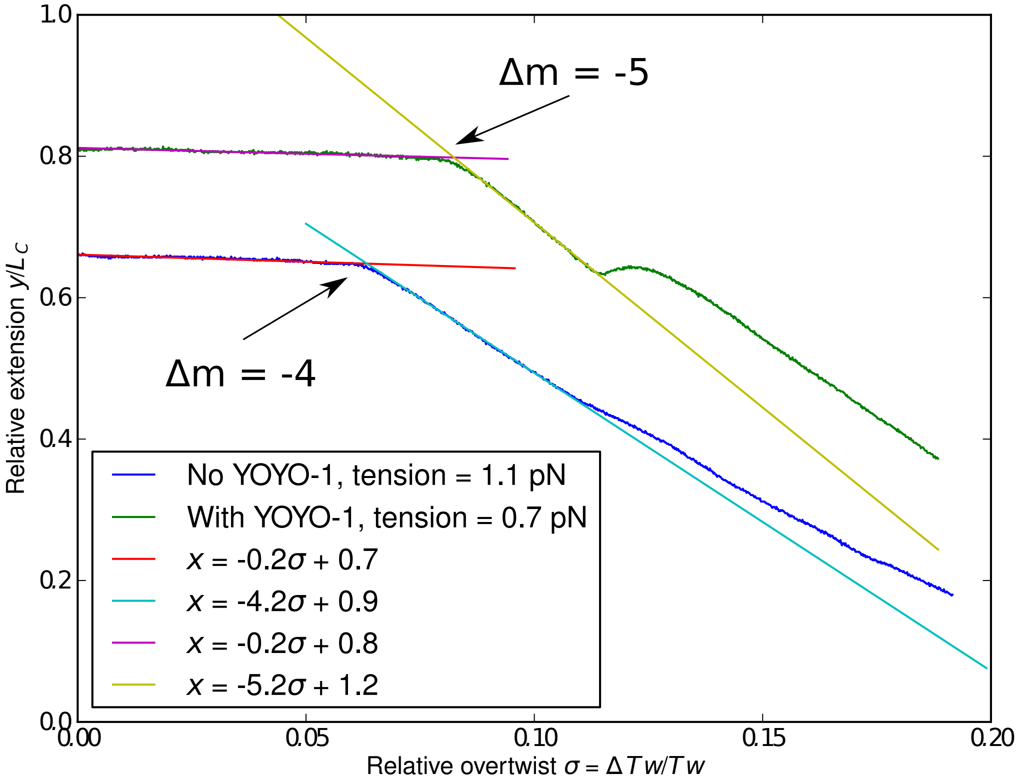


Figure 3.13: Comparison of the gradients before and after the buckling points for DNA with and without YOYO-1 for a single DNA/YOYO-1 complex. The gradient changes are marked with arrows and values given as $\Delta m$. Both systems exhibit similar behaviour in the low $\sigma$ regime. The buckling transition for DNA with YOYO-1 is more violent than for bare DNA by a factor of 1.25, indicating a sudden and more dramatic conformational change. Buckling points found as in [145].

The higher tension in the case without YOYO-1 present could account for the difference in buckling, as prior work has shown that at these low applied forces a small change can result in greatly changed behaviour [92]. However as the change in gradient in Figure 3.12 is higher than that in Figure 3.6 it is likely that the YOYO-1 intercalation is having a mechanical effect on the DNA construct.

Greater torsional stability due to the intercalated YOYO-1 could provide an explanation. If the structure is stabilised by the YOYO-1 but the support is removed – i.e. if a YOYO-1 molecule unbinds – the built up twist would be relaxed at the no-longer stabilised site. An unbinding event of YOYO-1 could therefore lead to a more dramatic length shortening than seen without YOYO-1 present at all: an unbinding of YOYO-1 in a small region could act as a nucleus for plectoneme formation. However the plectoneme forms extremely rapidly at the unbinding site as all the torsional stress in the system is concentrated there, so in a short amount of time a large amount of length is lost, and therefore the gradient change is dramatically steeper than seen in other systems. Fluorescence imaging of these systems will confirm or otherwise this speculative model of the system's behaviour.

### 3.2.3 Fluorescence Imaging with YOYO-1

Fluorescent images of DNA tethers were also obtained for both overtwisted and relaxed DNA. Figure 3.14 shows the effect of illumination on an overtwisted DNA tether, imaged with no free radical scavenging system present. The DNA, which had previously been stable at that extension, suddenly begins to supercoil further and shorten dramatically before snapping. This resulted in such a force being applied that the bead was pulled from the optical trap.

This fluorescence micrograph is more evidence of the structural effects described in the preceding sub-Section. The YOYO-1 being illuminated has directly resulted in the DNA creating more plectonemes and exerting more force, implying the the YOYO-1 was providing a structural stability that was lost when the dye photobleached and/or dissociated from the DNA. The dissociation leading to an abrupt, forceful structural transition in which the extension dramatically reduces would explain the violent behaviour of the DNA in Figure 3.10, except that in that case the dissociation would appear to be driven by purely mechanical rather than photophysical means.

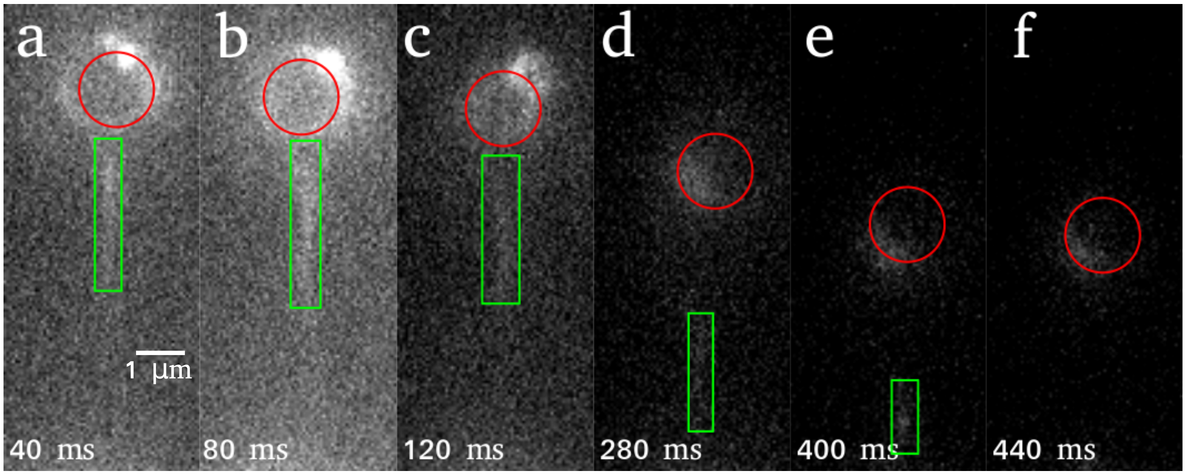


Figure 3.14: Fluorescence image of a tether snapping under illumination. The DNA had been overtwisted and under illumination suddenly supercoiled before snapping in e). The red circles indicate the magnetic bead position and the green boxes encapsulate the visible DNA.

While interesting, this snapping behaviour is not what is required of a fluorescence microscope. The power was thus lowered using a neutral density filter to 1 mW, and DNA was imaged in the full oxygen scavenging buffer. The result of this can be seen in Figure 3.15. The DNA tether is visible for over 200 frames, a total time of nearly 1 s. However, using a low imaging power has drawbacks. Firstly, the result is that the DNA appears less sharp, and the image is more blurry. Second, the high power is needed to force the YOYO-1 dyes into a dark state and allow them to photoblink for STORM reconstruction later. Using the lower power results in numerical analysis of the tether being very difficult if not impossible, while the lack of photoblinking prohibits super-resolution.

However, the two images together show that individual DNA tethers can be imaged using YOYO-1 in conjunction with the magneto-optical tweezers. This is clearly an area in which significant further optimisation will be needed to push the project on and achieve the results that this microscope is capable of, but the images show that such a possibility does exist.

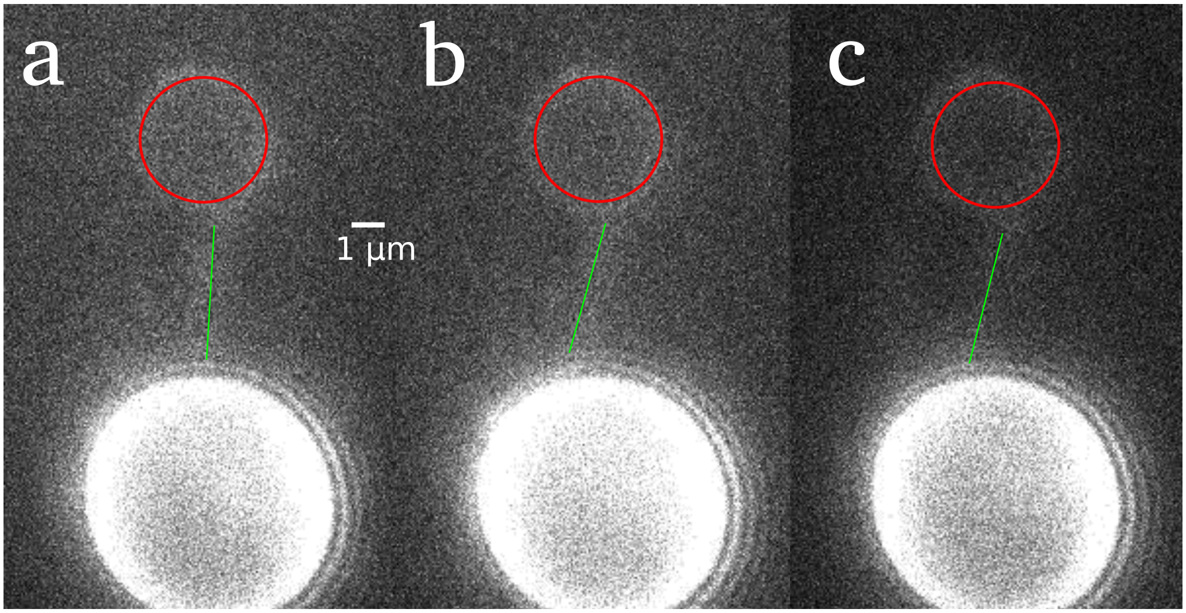


Figure 3.15: DNA tether imaged with 1 mW excitation laser power measured at 491 nm in oxygen scavenging buffer. Under these conditions, the DNA did not snap and was visible for around 1 s. While the increased longevity of the DNA is good, the tether is out of focus which, combined with the resulting loss in image sharpness due to the low imaging power, makes quantitative analysis of the tether largely impossible. Red circles indicate bead position while the green line indicates the tethered DNA.

## 3.3 Summary

The magneto-optical tweezers have been shown in this Chapter to perform experiments as desired. Force extension curves were generated with both labelled and unlabelled DNA. Both showed similar persistence lengths of 30-70 nm, which is in accordance with values found in the literature. However, the spread of persistence lengths was high, and more repeats are needed to build up a reliable distribution of persistence length for the tether under these conditions. Future work should begin by performing the same suite of experiments in the full imaging buffer as well as a less complex buffer which can be directly compared with published work. The position of the bead prior to a force-extension curve being generated should also be automatically changed to ensure there is minimal force in the direction perpendicular to the stage oscillations. This would reduce the complexity of analysis and increase robustness.

However, the consistency of results is good, and the persistence length did not change upon DNA binding in agreement with previous studies [84, 85]. The fitted contour length was shown to be similar to those found in the case without YOYO-1, in contradiction with those studies, however. The effect of the imaging buffer on the bare DNA again needs to be made clear in future as the fitted contour length was consistently $\sim 1\,\mu$m higher than an estimated expected value based on the structural parameters of B-DNA.

Holding tethers at constant force, different twists were applied. The DNA responded in the expected way: for higher forces ($F \geq 0.7\,$pN), undertwisting did not shorten the tethers but overtwisting did, while for lower forces around 0.55 pN either perturbation did. This behaviour was in good agreement with previous work, with the threshold occurring between 0.55 and 0.7 pN in comparison to the 0.6 pN to 0.8 pN demonstrated in [92]. Future work should apply lower forces both with and without YOYO-1 present to investigate any mechanical differences due to the dye binding. In these assays, the specific experimental conditions may have had a significant effect: temperature, salinity, and interactions with other molecules will change DNA's mechanical properties, and so the experiment should be repeated with a previously used buffer with fewer components. For example, [144] simply used 10 mM PB buffer.

Upon overtwisting and untwisting, and undertwisting and retwisting, both labelled and unlabelled DNA showed significant hysteresis-like behaviour, with the DNA taking different routes to and from its perturbed steady state. Whether hysteresis is expected in these DNA systems is unclear in the literature, with at least two papers published showing conflicting results [99, 145]. Again in future using a different buffer and repeating the experiments further will shed light on this complex issue. It is likely that the twisting rate plays a significant part, and future repeats should be done both at 2 Hz but also at a significantly lower rate of perturbation, with more equilibration steps built in – for example, apply $\sigma = 0.01$ at 0.1 Hz and equilibrate before applying more torsion.

The twist-extension plots also gave insight into the mechanical properties of the DNA/YOYO-1 complex. The buckling transition for a YOYO-1 labelled DNA tether were more severe than that for an unlabelled tether. Through buckling analysis done by fitting straight lines to the different twist-extension domains, it was shown that the buckling transition with YOYO-1 present resulted in a change of gradient approximately 1.25× higher than that observed for DNA on its own. The buckling was also more severe for the labelled YOYO-1 tether than for an unlabelled DNA tether under lower applied force, suggesting a mechanical effect of the YOYO-1. However, the forces applied in each case were similar but not equal. Further work may be needed to assess whether this effect is due to applied tension or is a property of the DNA/YOYO-1 complex.

The step in the gradient is indicative of a significant conformational change and is likely cause by a transition to a writhe-dominated system triggered by mechanical unbinding of the YOYO-1, a hypothesis which is bolstered by the fluorescence evidence in Figure 3.14. The fluorescence images show an overtwisted DNA supercoiling as a result of laser illumination, indicating that there is a photophysical effect. Bleached YOYO-1 dissociating from the DNA could lead to perturbations previously stabilised by the YOYO-1 to manifest in the now unlabelled areas, forming a condensation nucleus for supercoiling and plectoneme formation. As with the brightfield images, more repetition is needed to understand this complex interplay.

The fluorescence images using lower laser power demonstrate the possibility of performing these experiments with both fluorescence and mechanical manipulation, but it is in need of significant further optimisation. With low power, images may be obtained but quantitative analysis is difficult and super-resolution impossible. However, the possibility exists, and with further work it will become a powerful tool for interrogating nucleic acid interactions.

Overall, the data in this chapter shows that the microscope is now capable of generating tethers and performing complex, multistage experiments, and obtaining biologically relevant data in agreement with previous studies. The minor disagreements with existing literature are likely to be as a result of experimental conditions as opposed to a genuine systematic error in either the hardware or the analysis, and by performing control experiments for comparison, insight into the factors influencing the results will be gained. The microscope, with the methods outlined here, represents a powerful new vector for exploring the molecular workings of life.

# Chapter 4

# A Brief Guide to Molecular Simulation

In this chapter, the principles and mathematical formulation of forcefield-based molecular dynamics will be presented. Two codes were used during this project: oxDNA [31], a coarse-grained nucleic acid simulator; and Amber [146], an all-atom general purpose biomolecular simulation package. For each program, the underlying assumptions and models will be made clear, and their limitations discussed. Methods to apply force and torque to DNA in each regime will be compared both to each other and the aims of the project. The application of molecular dynamics simulations to experimental work will also be discussed as was done in a recent review paper [42], and the specific computational strategies employed in this work will be justified.

## 4.1 Why Perform Multiscale Modelling on DNA?

Both coarse-grained and all-atom simulations of biopolymers have been used for some time. Generally, coarse-grained and atomistic simulations have occupied different length and time domains: at present it is possible to simulate dozens to hundreds of DNA base pairs over hundreds of nanoseconds to microseconds with atomistic modelling, and thousands of base pairs over timescales ranging to seconds with coarse-grained simulation. The power of each of these approaches combines to more than the sum of the individual parts, and correlating multiscale simulations of DNA performed under the same conditions and with the same methodology will allow investigation of atomic-level details such as hyrodgen bonding, while also examining the broader picture of a significant contour of DNA and assessing the localisation of perturbations.

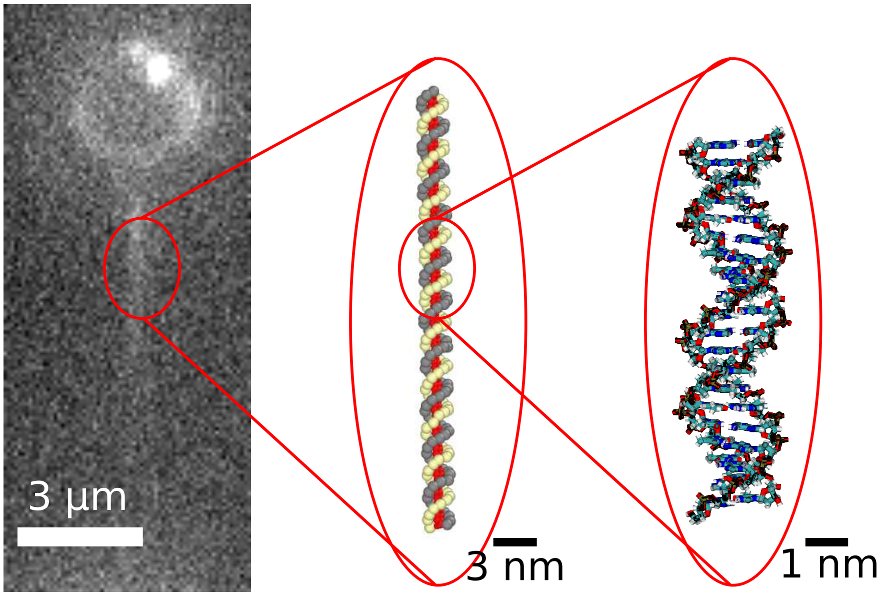


Figure 4.1: Magnifying a DNA experiment with successive levels of molecular simulation. Left to right: a fluorescence micrograph; oxDNA coarse-graining; the Amber atomistic model

Coarse-grained simulation can now be performed on fragments hundreds or thousands of base pairs long: nearly large enough to be compared directly to experiment, and doing this both validates the coarse-grained simulation trajectory and crucially acts as a link between experiment and atomic detail of DNA structures. By exploiting each tool to extract the maximum of information, a comprehensive picture of a system of interest can be observed and regions of interest selected for detailed analysis similar to the magnification in Figure 4.1.

## 4.2 The Models

### 4.2.1 Amber

The Amber [146] model of the molecular world is a relatively intuitive one: each atom or ion in the physical system is represented by a particle in the simulation. During the simulation, interaction energies are calculated as a function of distance, and the position and velocity of each particle is updated and the trajectory propogated through time.

Thermal effects are introduced using a stochastic noise function, and the overall thermal energy of the system is maintained by a thermostat – most commonly the Langevin thermostat. The random noise element combined with the thermostat can be shown to produce an ergodic system. A system is ergodic if for a long enough simulation the system explores all possible microstates with a probability proportional to the energy of each individual state. States with the same energy are therefore sampled with the same probability, and higher energy states are sampled proportionally less. For this to be true, the system needs to be in thermodynamic equilibrium. Therefore in order to get good sampling, it is common to simulate a system until its potential and kinetic energies, temperature, and pressure of the system have stopped changing but are instead fluctuating around a constant mean value. The production molecular dynamics is taken after this point to ensure good sampling. This overall process is known as equilibration.

In the atomistic model, the interactions between each pair of atoms, whether covalent, electrostatic, van der Waals, or hydrogen bonds, are specified by the forcefield where for each pair of atoms the constants used in the calculation of the potential energies are defined. For example, for a carbon-oxygen single covalent bond approximated as

$$\begin{aligned} U &= \frac{1}{2}k(x - x_0)^2 \\ \vec{F} &= -\vec{\nabla}U \end{aligned}$$

where $U$ is the potential energy of the bond, $x$ the distance separating the atoms, $x_0$ the equilibrium separation for the atoms and $k$ the spring constant, $x_0$ and $k$ would be defined by the forefield specification.

This raises an obvious question: how should these values be determined? Often an *ab initio* quantum chemistry calculation is used, as a geometry optimisation will trivially give $x_0$ and calculations at different separations will give forces to which the constant parameters may be fitted.

There are more subtelties to this approximation. For example, in a covalent bond, where are the charges? For a covalent bond the common charge can be thought of as distributed between the atoms depending on their charge and valence, but it is the charges which are used for electrostatic interactions with other atoms. Force fields do not allow for electrostatic interactions with a semi-fictional distribution. Instead, partial charges are distributed among the bond's members according to distributions calculated by all-electron quantum chemistry calculations. Thus in a notional bond if one atom has atomic number $Z = 10$ and the other one $Z = 20$ and the charge is distributed 70%:30% respectively, the new charges would become $Z = 10.7$ and $Z = 20.3$.

Single covalent bonds may be rotated around but if there are multiple bonds they exert torsion when twisted. As a consequence, there is an equilibrium bond angle and an effective spring constant, again worked out with quantum chemistry and again the domain of the forcefield during an atomistic simulation. Similarly, the angles between atoms' electronic orbitals can exert a force due to twisting, and this again an equilibrium position and a spring constant.

Putting together these different sets of contributions, we could say that a forcefield is overall

$$U_{total} = U_{covalent} + U_{electroststic} + U_{bondAngles} + U_{nonbonded} \tag{4.1}$$

Taking each pairwise interaction in turn, for two atoms $i$ and $j$,

$$U_{ij,covalent} = k_{ij,x} \left( x_{ij} - x_{ij,0} \right)^2 \tag{4.2}$$

where $k_{ij,x}$ is a spring constant dependent on the value of the displacement $x_{ij}$ i.e. $k_{ij,x} = K_{ij} \left( x_{ij} \right)$. For electrostatics the familiar equation

$$U_{ij,electrostatic} = \frac{q_i q_j}{\epsilon x_{ij}} \tag{4.3}$$

is used, where $q_i$ and $q_j$ are the modified charges on each particle as discussed above and $\epsilon$ is the permittivity of the medium (absorbing factors of $4\pi$). The nonbonded interactions are simply modelled as a Lennard-Jones potential [147]

$$U_{ij,nonbonded} = \frac{A_{ij}}{x_{ij}^{12}} - \frac{B_{ij}}{x_{ij}^{6}} \tag{4.4}$$

where $A_{ij}$ and $B_{ij}$ are factors of proportionality which depend on the atom types.

For the angular contributions, the energy contribution is defined as

$$U_{bondAngles} = \sum_{angles} K_\theta \left( \theta - \theta_0 \right)^2 + \sum_{dihedrals} \frac{V_n}{2} \left[ 1 - \cos \left( n\phi - \gamma \right) \right] \tag{4.5}$$

where $K_\theta$ and $V_n$ are constants of proportionality, and $\theta, \phi$ and $\gamma$ define bond angles and torsions.

Rather than calculate the electrostatic potentials for every pair of atoms in a system, a method called Particle Mesh Ewald (PME) [148] making use of Ewald summation was developed and is used to calculate electrostatic forces beyond a user-specified cut-off distance. In the simulations making up this work however the systems were small and computationally cheap enough that PME method was not used. Instead, the electrostatic cutoff was set to an arbitrarily large value ensuring that all electrostatic interactions were treated explicitly.

In this work, the Amber forcefield *bsc1* [149] was selected. This is based on the very successful *bsc0* [150] modifications to the Amber *parm99* [151] forcefield, and has been shown to produce good results over a range of DNA simulation conditions [152].

### 4.2.2 oxDNA

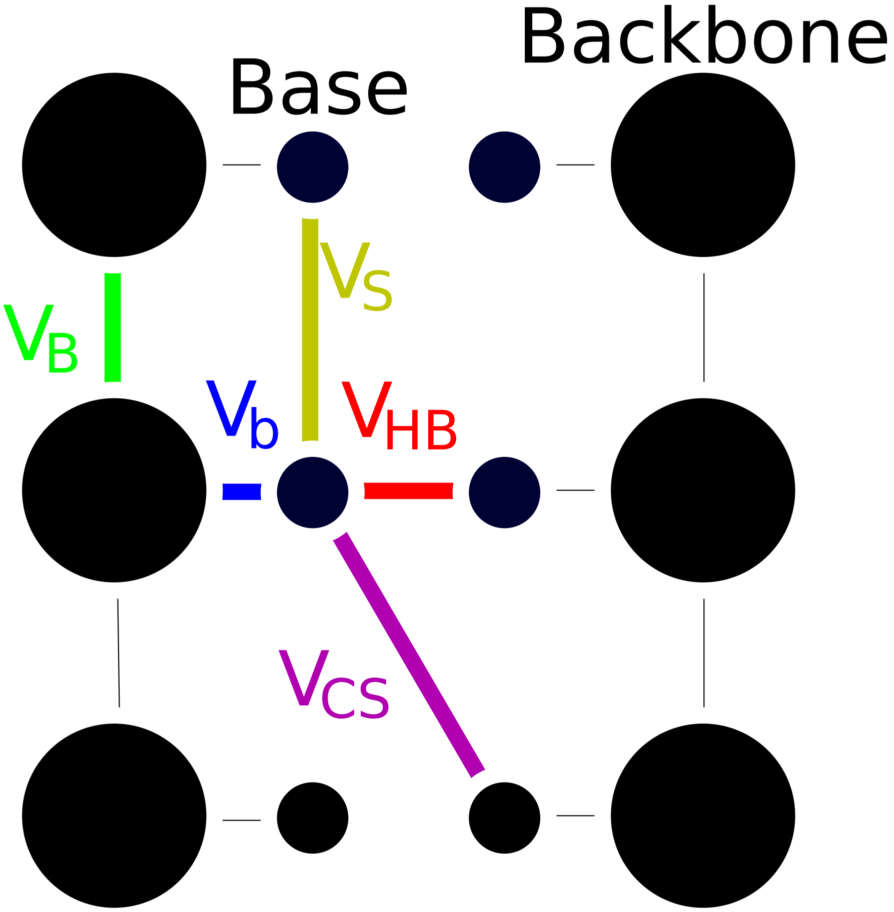


Figure 4.2: The oxDNA model with interactions for the left middle base indicated

It is remarkable that the oxDNA model of nucleic acids should have become so successful and simulated DNA and RNA so accurately considering the amount of detail it neglects. The oxDNA model states that nucleotides can be reduced to two components: the backbone/sugar ring and the base itself. Each nucleotide is therefore composed of only two particles, each base pair represented by four spheres. A pictoral representation of this stripped-down model is in Figure 4.2.

In this model there are just sixteen interactions per base pair with no base differentiation. Though oxDNA has subsequently got support for base sequence, the early version and much of the groundbreaking work utilised only one "average" nucleotide. The nucleotide interacts with its neighbours according to the backbone bonds, and the bases interact at a longer range with Lennard-Jones style potentials. These potentials are straightforward and computationally cheap, and to cut computational cost yet further the nonbonded interactions extend only to nearest neighbours. One base pair therefore interacts with its neighbours as shown in Figure 4.2. In total, one base pair has three bonded and five nonbonded sources of potential energy, and all particles additionally interact via a weak Lennard-Jones potential. Specifically, these are [153]:

- $V_B$: the backbone-backbone interaction
- $V_b$: the backbone-base interaction
- $V_{HB}$: the complementary base-base hydrogen bond interaction
- $V_S$: the stacking interaction between subsequent bases on the same strand
- $V_{CS}$: the cross-stacking interaction, which exists between neighbour bases on opposite strands, i.e. between base $i$ on one strand and bases $i \pm 1$ on the other
- $V_{LJ}$: weak repulsion between particles

In this model, electrostatic interactions are excluded and treated only with excluded-volume terms, $V_{EV}$, which will not be discussed here as their technical depth falls outside the scope of this thesis. Both $V_{HB}$ and $V_S$ have the same functional form ($V_M$) but are differently parametrised [154].

Again, the forcefield in oxDNA defines the potential energy due to each interaction type. These are [154]

$$\begin{aligned} V_B =& -\frac{\epsilon}{2}\ln\left(1-\frac{(r-r_0)^2}{\Delta^2}\right) \\ V_M =& \epsilon\left(1-\exp\left(-a\left(r-r_0\right)\right)\right)^2 \\ V_{CS} =& \frac{k}{2}\left(r-r_0\right)^2 \\ V_{LJ} =& 4\epsilon\left(\left(\frac{\sigma}{r}\right)^{12}-\left(\frac{\sigma}{r}\right)^{6}\right) \end{aligned} \tag{4.6}$$

where $\epsilon, r_0, \Delta, a, k$, and $\sigma$ are fitting parameters for each interaction type.

Given that a single nucleotide in Amber has above fifty atoms, all of which have multiple interactions, it is clear to see how dramatic the speed difference between the codes can be.

## 4.3 The Langevin Thermostat

Of key importance to a molecular dynamics simulation is maintaining a constant biologically relevant temperature, and in this work the Langevin thermostat was used for this. The Langevin thermostat is a conceptually straightforward scheme which includes Brownian motion in the form of a stochastic force applied to each atom at each timestep. There are various implementations but the Amber implementation based on that published by Loncharich, Brooks, and Pastor [155] which is briefly summarised below.

The Langevin equation for a system evolving due to applied force, frictional force, and random noise is given by

$$m\frac{\mathrm{d}^2x}{\mathrm{d}t^2} = F(t) - \xi\frac{\mathrm{d}x}{\mathrm{d}t} + R(t) \tag{4.7}$$

where $m$ is the particle mass, $\frac{\mathrm{d}^2x}{\mathrm{d}t^2}$ is the particle's acceleration, $F(t)$ is the force exerted on the particle by the system, $\xi\frac{dx}{dt}$ gives the strength and direction of the resistive force acting against particle movement, and $R(t)$ is the stochastic Brownian fluctuation which is taken from a Gaussian distribution centred on 0 with variance

$$\sigma^2 = \frac{2m\gamma k_B T}{\Delta_t} \tag{4.8}$$

where $\gamma = \frac{\xi}{m}$, $k_B$ is the Boltzmann constant, $T$ is the temperature and $\Delta_t$ is the timestep.

Numerical integration of this Langevin equation is done in two steps using the leapfrog algorithm [155]. This system of equations integrates between $t_n$ and $t_{n+1}$ in two stages, integrating the velocity first to the midpoint $t_{n+\frac{1}{2}}$ by approximating the acceleration, using this new velocity to integrate the atomic position to $t_{n+1}$, and then finishing the velocity integral to $t_{n+1}$. Mathematically [155],

$$\begin{aligned}
v_{n+\frac{1}{2}} &= v_{n-\frac{1}{2}} + (x_n - x_{n-1})\frac{1-\frac{1}{2}\gamma\Delta_t}{1+\frac{1}{2}\gamma\Delta_t} + \left(\frac{\Delta_t^2}{m}\right)\cdot\left(\frac{F_n + R_n}{1+\frac{1}{2}\gamma\Delta_t}\right) && (4.9)\\
x_{n+1} &= x_n + \Delta_t \cdot v_{n+\frac{1}{2}} && (4.10)\\
v_{n+1} &= \sqrt{1+\frac{1}{2}\gamma\Delta_t}\left(\frac{v_{n-\frac{1}{2}} + v_{n+\frac{1}{2}}}{2\Delta_t}\right) && (4.11)
\end{aligned}$$

where again $\gamma = \frac{\xi}{m}$, $\Delta_t$ is the timestep of integration, $x$ and $v$ represent position and velocity respectively, $n$ is the timestep, $F_n$ and $R_n$ are the total force and Brownian force on the atom at timestep $n$, and $m$ is the atomic mass.

The thermostat itself works through the interplay of $T$ and $\xi$. Changing these values results in atoms on average gaining additional velocity, losing velocity to drag, or on average keeping the same velocity, resulting in a system with increasing, decreasing, or constant temperature respectively. The stochastic force $F_n$ means the system is not completely deterministic and makes sure the system satisfies the ergodic condition.

## 4.4 Implicit Solvation

Modelling water is a computational challenge for multiple reasons. It is a polar molecule with unusual properties at temperatures around its freezing point, but more difficult than that in a classical molecular dynamics simulation is that there are simply so many water molecules to model in a biological system. The solvation shell of DNA has been extensively studied with molecular dynamics, and so now the water is itself not of interest – it is there to provide the correct conditions for simulation. Removing the water molecules altogether is therefore an attractive proposition as cheaper computation increases the scale of possible simulations, especially since the lack of explicit water lowers the frictional force experienced by biomolecules. As a result, the system's exploration of configuration space takes less simulation time.

However, it is important to make sure all the biomolecules in the simulation are properly parameterised for non-explicit water as those intended for work only in explicit solvent would give incorrect dynamics. For this reason, small-molecule interactions with DNA are in general best performed with explicit solvation, while studies of DNA on its own, for example minicircle topology, have been successful in implicit solvation [60, 54, 55]. In this work, the use of implicit solvation is well-justified and makes possible greater exploration of the DNA dynamics under torque and force than would otherwise be possible.

### 4.4.1 The Generalised Born Solvation Model

A simple implicit solvation method would be to change the dielectric constant governing electrostatic interactions as a function of the distance between the charges considered, i.e. $\epsilon(r)$. This would give effective screening of charge-charge interactions, which is one of the effects of water, but it is in practice inaccurate over long timescales and much better methods have been developed. In this work, the commonly used Generalised Born (GB) solvation system [156, 157] is used. Similarly to a simple dielectric scheme, the overall effect of the GB model is to approximate the effects of water as a function of distance. However, it has a key difference: the solute and solvent are treated as separate regions with different dielectric constants – the radius of this boundary must be parameterised for each atomic species and is called the effective Born radius. This is an important physical quantity and the system is extremely sensitive to it so values must be chosen with care [158]. The Poisson equation which governs the contribution to the free energy due to electrostatics has an exact mathematical form [158] but in the computational implementation is approximated as [156]

$$\Delta G(\text{kCal/Mol}) = -166\left(1-\frac{1}{\epsilon}\right)\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{q_i q_j}{f_{GB}} \tag{4.12}$$

where $\epsilon$ is the unitless relative permittivity, $N$ is the number of atoms in the system, $q_i$ and $q_j$ are the charges on the two atoms under consideration (measured in electron charges), and

$$f_{GB} = \sqrt{r_{ij}^2 + \alpha_i\alpha_j e^{-D}} \tag{4.13}$$

$$D = \left(\frac{r_{ij}}{2\alpha_i\alpha_j}\right)^2 \tag{4.14}$$

where $\alpha_i$ is the effective Born radius in Å for atom $i$ and $r_{ij}$ (Å) is the distance between the two atoms. Here the approximation used that the solvent can be treated simply as a continuum, and the change in free energy due to the solvent is directly proportional to the solvent-accessible area. When the distance between the particles goes to infinity (essentially, for a single particle is on its own), the change in free energy due to the solvent is reduced to a constant of -166 kCal/Mol.

These parameters clearly include species-specific effects as well as different interaction strengths between different species. When it was announced, this fit known data very well [156] and continues be refined.

#### 4.4.1.1 Ionic Strength

In experiment it has been shown frequently that biomolecular structure, interactions, and dynamics are strongly dependent on the salt concentration which can screen charges and thus allow charged regions to come closer together than would ordinarily be allowed. In the Amber GB solvation model, the salt concentration is varied by inclusion of a

Debye-Huckel correction, so that the change in free energy becomes [159]

$$\Delta G = -\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{q_i q_j}{f_{GB}}\left\{1 - \frac{e^{-\kappa f_{GB}}}{\epsilon}\right\} \tag{4.15}$$

where $\kappa$ is the Debye-Huckel screening parameter. Setting $\kappa$ to a nonzero value results in varying the salt concentration.

### 4.4.2 Solvation in oxDNA

In oxDNA the picture is very different from the all-atom philosophy of Amber. The aggregation of atoms into backbone and base pseudoparticles means that the detail needed to capture electrostatic interactions is much reduced. Indeed, the initial model did not account for electrostatic effects at all. Instead, excluded volumes around the centres of the simulation particles – regions that other particles were not allowed to enter and were repelled from – gave rise to interations which were effectively electrostatic in nature. The hard-coded values in the excluded volume calculation [154] meant that varying the strength of the interaction was not possible, and the overall effect was to simulate the DNA in a high-salt (i.e. strongly electrostatically screened) medium – approximately 600 mM salt equivalent [154]. However, later updates allowed variation of the excluded volume parameter, changing the size of the excluded region. This allows the particles closer together or forces them to be further apart, mimicking the screening effect of salt so that simulation of different salt concentrations are now possible.

## 4.5 Applying External Force and Torque

There are many means to apply force in molecular dynamics simulations, which can broadly be split into one of two categories: application of defined forces to defined atoms, or stressing the simulation cell such that a force is generated on the system within. In this work, only the former category was utilised. Within this overall philosophy of targeting individual atoms, there are two primary techniques:

**Positional or angular restraints** which set up a potential well around given positions for given atoms, and so confine the atoms to remain in that location. Variation of the strength of the potential wells allows the atoms to be fixed confined more strongly or weakly, and thus affects how far from the equilibrium position an atom can stray. This is analogous to affixing an atom to a spring of given stiffness. The equilibrium position of the trap may in some software be allowed to move with time and thus apply a stretching or twisting force. It is important to note here that positional *restraints* are distinct from positional *constraints*. A constraint implies that the atom is not allowed ever to move, resulting in a completely hard wall, and is therefore essentially a positional restraint with infinite force constants. Care must be taken with both restraints and constraints as a fixed atom can introduce extraneous energy into the system – if an atom moves too close to it there will be an extremely high repulsive force. At the cost of higher computational expense, for this reason a shorter timestep should be used when dealing with restrained or constrained systems.

**Specific force application** allows a user to select atoms and specify a force magnitude and direction, which may vary with space and time. A conceptual example of this may be applying an electric field to selected atoms.

Positional restraints were selected as the most appropriate for this work, for numerous reasons. First, the Amber package most easily allows application of this kind of force. Secondly, in the experimental system the beads are held a fixed distance apart by the force clamp, and although this is in fact an application of force, it has the effect of fixing end-to-end length. Thirdly, fixing the positions of the atoms prevents the DNA from relaxing applied torsion, and it is in this dual application of force and torque which the magneto-optical tweezers are designed for. Allowing the DNA to rotationally relax would not be comparable to experimental conditions.

Amber does not allow the user to move applied positional restraints during a simulation. There are two distinct ways to stretch the DNA: either the terminal bases can be moved a distance away from their starting positions, then fixed in this new position; or alternatively a harmonic positional trap can be moved and the atoms drawn towards it. It is most straightforward to implement the former method, and as it looks particularly at the equilibrium state of a stretched molecule rather than its transition to an equilibrium state, this technique is well suited to comparison with experiment.

To stretch DNA according to this method, the canonical structure is built, and one of the terminal base pairs moved by a distance $\Delta z$ along the helical axis in a stretching direction by modification of co-ordinates in the structure file. Harsh (500 kcal$\cdot$mol$^{-1}\cdot$Å$^{-1}$ in Amber) positional restraints are then applied to all the atoms in the terminal nucleotides to keep them in their new positions. A stretch of $\Delta z$ is thus applied and after energy minimisation the production simulation can proceed. The final frame of this simulation is then taken as the starting structure for the second round of stretching, where once again one of the terminal base pairs is moved by $\Delta z$, fixed in place, and the system simulated. This procedure is repeated the desired number of times to simulate stretching the DNA to a desired amount beyond its relaxed contour length. A graphical representation of one such cycle is given in Figure 4.3.

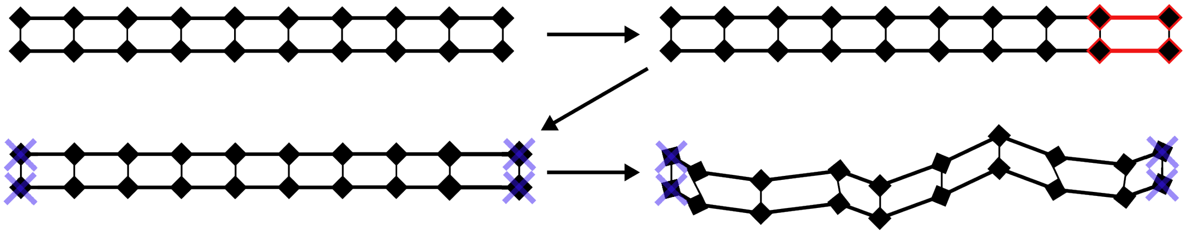


Figure 4.3: The stretching procedure for simulations in this work. Clockwise from top left: the starting structre; the starting structure with one of the terminal base pairs moved by $\Delta z$ (indicated in red); positional restraints applied to all terminal base pairs (indicated with blue crosses); the system is simulated

oxDNA, by contrast, allows the user to move a trap at a given rate in a given direction while the possibility remains also to replicate the Amber procedure. To get the most comparability between oxDNA and Amber, the stretching routines should be identical, and so the procedure designed for Amber simulations (Figure 4.3) was used also for oxDNA.

Application of positional restraints also makes it easy to vary the applied torsion. By applying a rotation matrix to the atomic positions of the terminal base pair in the DNA fragment the twist in the fragment is varied by $\Delta\theta$ in either an undertwisting overtwisting sense, which the applied positional restraints maintain. However, in supercoiling experiments the change in twist is not usually described by the change in angle itself. Instead, the parameter $\sigma$ is defined as

$$\sigma = \frac{\Delta Tw}{Tw_0} \tag{4.16}$$

where $Tw_0$ is the total twist in the relaxed system (i.e. for two full helical pitches $Tw_0 = 720$) and $\Delta Tw = Tw_0 - Tw$ is the angular perturbation applied. Common experimental values for $\sigma$ are between $-0.05$ and $0.05$ [92] though $\sigma$ values an order of magnitude higher than this have been used. In order not to over-perturb the DNA, overstretch bonds, and generate enormous energies and forces in the starting structures, large rotational perturbations are performed as a series of small rotations followed by rigorous structure minimisations (see Section 4.6) in a manner analogous to the stretching protocol outlined above.

## 4.6 Minimising the Potential Energy of the System

Building a system for simulation is generally done using experimental models – for example, Amber's `nab` utility builds DNA helices using user-specified sequence and the canonical parameters taken from fibre diffraction. However the structures generated are not guaranteed to be free from problems such as stretched bonds or atoms too close to each other, and this is especially true if the user has modified the starting structures, for example by moving a base pair. Before finite-temperature production molecular dynamics can take place it is therefore necessary first to minimise the system to relax over-stretched bonds and allow steric clashes to be resolved in a safe way.

To minimise the potential energy of a structure, the forces on each atom are calculated at 0K – i.e. without their having velocity and without thermal Brownian effects being present. The atoms are then moved a small distance towards their minimum energy position, a direction which is determined in one of two primary ways:

- *Steepest descent,* in which the atoms are moved in the direction of the overall force they are subject to

- *Conjugate gradient*, in which the system of positions and forces is expressed as a linear algebra optimisation problem to be solved, resulting in moving the atoms not in the direction of the gradient but in a direction conjugate to it. This can demand significant computational resources and requires the problem to be expressed as sparse matrices. The conjugate gradient method can be unstable far from a solution but often requires fewer iterations than steepest descent.

Amber makes use of both schemes, by default performing steepest descent followed by conjugate gradient minimisation. oxDNA meanwhile only performs steepest descent optimisation.

In each case, the atoms are moved incrementally and therefore this is an iterative instead of one-shot procedure, and in Amber the two methods are typically used in conjunction with one another. The steepest descent method is used for a certain number of iterations to get the system close to its minimised structure, and the conjugate gradient method then takes over for refinement. The minimisation continues until either the total energy change of the system between steps $n$ and $n + 1$ is below a threshold value, or until a fixed number of steps has been completed. Condensed matter and quantum chemistry simulation software generally use a convergence parameter to test $\Delta E$ against, where $\Delta E$ is the change in potential energy between successive minimisation steps. Molecular dynamics packages (which require less precision) usually have users specify a number of iterations to be performed.

In this work, significant perturbations are applied to the DNA. It is therefore necessary to minimise the structure well. The conjugate gradient method in these conditions (stretched, rotated base pairs) was found to be particularly unstable and consequently the overall strategy relied heavily on steepest descent minimisation. To ensure that the energies were properly minimised despite the perturbations, for Amber the total energy was plotted as a function of minimisation steps for both rotated and stretched DNA. oxDNA does not have an option for outputting minimisation energies, so instead the total energy was examined for a simulation stage, and the quality of the minimisation inferred. The results of this minimisation will be presented in Chapter 5.

## 4.7 DNA Structures and Production MD

Having built the structures, rotated base pairs if appropriate, applied a displacement to one terminal base pair, and minimised the system thoroughly, production molecular dynamics can be performed. In both Amber and oxDNA the systems were held via the Langevin thermostat at 300K, with fixed ends. The total simulations were made up of 30 stretching events per rotation event, with $\Delta z = 1\,\mathring{A}$ and $\Delta\theta = 14°$ to give $\Delta\sigma \approx 0.017$. Four rotations in both the undertwisting and overtwisting senses were performed, as well as one stretching simulation with $\sigma = 0$. The structures were thus stretched by a maximum of 30Å for a total stretch of 1.3× and rotated by a maximum of $\sigma = \pm 0.068$.

Four structures were chosen: poly-d(A)·poly-d(T), poly-d(A)·poly-d(T), poly-d(C)·poly-d(G), and poly-d(CG)·poly-d(CG) to give a good account of the effects at the extremes of different densities of A, T, G, and C bases and their different stacking properties. These short fragments of repeating units are complex and wide-ranging enough to give meaningful predictions for experiments, but simplified and well-contained enough for tractable analysis. Repeating units of complementary base pairs should also show the effects of the stacking interactions in a dramatic way.

Overall there were thus 270 individual simulations per structure, each of which consisted of 500 ps of molecular dynamics, for a total simulation time of 135 ns per structure. With four structures, this led to 540 ns of simulation time in total for the 500 ps/Å/$\sigma$ simulations, a duration demanding significant computational resources and time, both in the simulation and analysis stages. Alongside these, two replicas were performed, which were identical except that the simulation time was 100 ps/Å/$\sigma$ for a total time of 135 ns per replica. The total simulated time in this project therefore was 810 ns in Amber.

In oxDNA, two sequences were chosen to qualitatively demonstrate differences or lack thereof between AT-rich and GC-rich DNA, specifically poly-d(A)·poly-d(T) and poly-d(C)·poly-d(G). 100 stretching events were simulated with each stretching event being of 0.21 nm so that the overall applied overstretch was 21 nm. Each stretching event was minimised with 100,000 steepest descent steps and simulated for 3 ns.

## 4.8 Analysis of Trajectories

Analysis of molecular dynamics trajectory is a complex undertaking. As mentioned in Chapter 1, there are various structural parameters which should normally be taken into account:

- Helical parameters including the overall contour length, minor and major groove depths and widths, and writhe
- Inter-base pair parameters, those concerned with the differences from one base pair to another. These typically include the twist, and rise between successive base pairs.
- Intra-base pair parameters, which determine how the bases making up the base pairs are oriented with respect to each other, including their pitch, opening and propellor

However, these parameters rely on the DNA fragment conforming relatively closely to canonical structures. Broken base pairs, non-canonical hydrogen bonds, and exotic base pair stacking can lead to misleading results. It is therefore necessary to analyse the structures in this thesis in a more fine-grained manner. Specifically, analysis of the individual hydrogen bonds which make up the Watson-Crick base pair interactions and the stacking interactions were undertaken, to quantify how these transient bonds may be creating novel structures. In this work, analysis and plotting of the canonical stacking and Watson-Crick hydrogen bonds was performed using software kindly provided by Dr Agnes Noy of the University of York. This software identifies canonical hydrogen bonding between donor and acceptor atoms based on the distances between them, and was subsequently modified by the author to also identify the presence of non-canonical hydrogen bonds which are not seen in less distorted structures. This dual analysis of canonical and non-canonical interactions was combined to produce detailed pictures of the conformational landscape experienced by DNA under torsion and tension.

## 4.9 Comparison to Experiment

These simulations are necessarily short both in length- and time-scale, and so relation to large-scale experiment is difficult. However, some computational methods employed make it somewhat easier. Firstly, the thorough structural minimisation should bring the structures into their most relaxed, steady state. This is what will be seen in the experiments due to the long time-scales that images are acquired over. Atomic analysis of the structures will also allow quantification of the disruption to canonical Watson-Crick base pairing. As YOYO-1 is an intercalator, a significantly disrupted region may see less YOYO-1 binding under experimental conditions. This is the key prediction that the analysed trajectories will give: the relative disruption to the DNA double helix due to different perturbations and sequences, which can be measured experimentally through quantification of fluorescence intensity. In this work, due to significant computational complexity, it was not possible to simulate a DNA/YOYO-1 complex under these conditions, which would have provided even stronger predictive power.

## 4.10 Summary

In this Chapter, the important concepts of classical molecular dynamics were explained and the key algorithms such as the leapfrog integration scheme and the Langevin thermostat, which make such simulations meaningful, were introduced. The means relevant to this work of applying external forces in simulations were discussed, and some limitations such as the need for a reduced timestep were commented on. Minimisation of the systems' energies in the context of perturbed structure generation was discussed and a suitable optimisation methodology was proposed. The parameters which analysis of trajectories relies upon were identified as were the external software packages used to perform such analysis, the chosen structures were introduced and justified, and the overall scheme for twisting and stretching DNA fragments was given.

# Chapter 5

# Simulating DNA

In this Chapter is presented the simulation work carried out to investigate the effect of sequence and applied twist on over-stretched DNA. The work in this Chapter is the sole work of the author with the exception of the analysis code to calculate canonical Watson-Crick hydrogen bond occupation, which was kindly written and provided by Dr Agnes Noy.

## 5.1 Simulation Design

As discussed in Chapter 4, a stretching force can be applied by a harmonic trap which moves during a simulation or by a trap which is fixed in space and moved between separate simulations. Amber does not support harmonic traps which move during a simulation, so a fixed trap system was designed. To make simulations as comparable as possible, the same procedure was followed done in oxDNA. The protocol carried out was as follows:

1. If applicable, rotate one terminal base pair by 14
2. Move the terminal base pair by altering co-ordinates in a simulation structure file (.pdb in Amber, .dat in oxDNA)
    (a) In Amber, construct simulation input files using tleap [146]
3. Apply harmonic traps to the terminal base pairs
4. Minimise the system as completely as possible
5. Run production molecular dynamics

6. Use the final frame of the production molecular dynamics as the input for the next stretching simulation

In general, molecular dynamics would be run after step 4 in order to allow the system to equilibrate and therefore to ensure good sampling of conformation space, and these results would be discarded prior to analysis of the production MD. However, in these simulations it was found not to be necessary. Although in the oxDNA simulations the structure was not completely equilibrated towards the end of the simulation (Figure 5.1), the distance to equilibration was very small. As the oxDNA simulations were used largely for qualitative reasons – to observe the localisation of perturbations – including an equilibration step would not materially affect the results. In Amber, meanwhile, the system was found to be almost instantly equilibrated, as will be shown in later sections. Only the first frame of each simulation stage was not in equilibrium, but this accounted for little enough of the simulation that its removal was unnecessary.

By repeating the procedure above, overall stretches of 30 Å for a 24 bp fragment (an increase in contour length of 39.6%) were performed in Amber, and stretches of 21.3 nm were performed on a 100bp fragment in oxDNA. In both regimes, molecular dynamics was performed at 300 K which was maintained by use of the Langevin thermostat. Timesteps were 1 fs in Amber with the DNA structure written to disk every 1 ps. Three replicas were performed for each configuration, one 500 ps per structure for a total simulation time of 15 ns per sequence per $\sigma$. $\sigma$ took values 0, $\pm 0.017$, $\pm 0.034$, $\pm 0.051$, and $\pm 0.068$. Two shorter replicas were performed with times of 100 ps per Å per $\sigma$ with structures and $\sigma$ the same as before.

oxDNA simulations had timesteps of 0.001 simulation units of time, equalling 3.03 fs. 100 stretching events were simulated with oxDNA and each event was then simulated for 100,000 steps for a total time of 303 ps. Two structures were simulated with oxDNA: poly-d(A)·poly-d(T) and poly-d(C)·poly-d(G) using the DNA2 forcefield [68] which specifies major and minor groove widths as well as different interaction strengths for individual bases. No difference in $\sigma$ was simulated due to the inability of oxDNA to properly simulate S-DNA in previous work [66] which implies that detailed structural information is obtainable only through all-atom simulations. Instead, oxDNA was used to assess how the structural changes induced by stretching are distributed along the contour length of the DNA. The two structures were chosen such that sequence-dependent effects could be readily seen. Due to the coarse nature of the oxDNA model, one CG rich construct and one AT rich construct was considered a good starting place – if significant differences in behaviour were seen then further simulations on other sequences could be performed. However, very little qualitative difference was seen, and as oxDNA was being used primarily as a qualitative indicator of the distribution of applied perturbations, no further sequences were investigated.

All simulations were performed using the NVT ensemble which preserves number of particles, the volume of the simulation cell, and the system's temperature, and both were simulated in an implicit solvent with effective 50 mM salt concentration. Amber simulations were performed with the Amber 16 version of `pmemd.CUDA` [146] on an nVidia GT 730, while the serial version of oxDNA 2.3 was used for oxDNA simulations [67, 68]. The oxDNA structures were visualised using Ovito [160, 161] while the Amber trajectories were visualised with VMD [162].

## 5.2 oxDNA Simulations

### 5.2.1 poly-d(A)·poly-d(T)

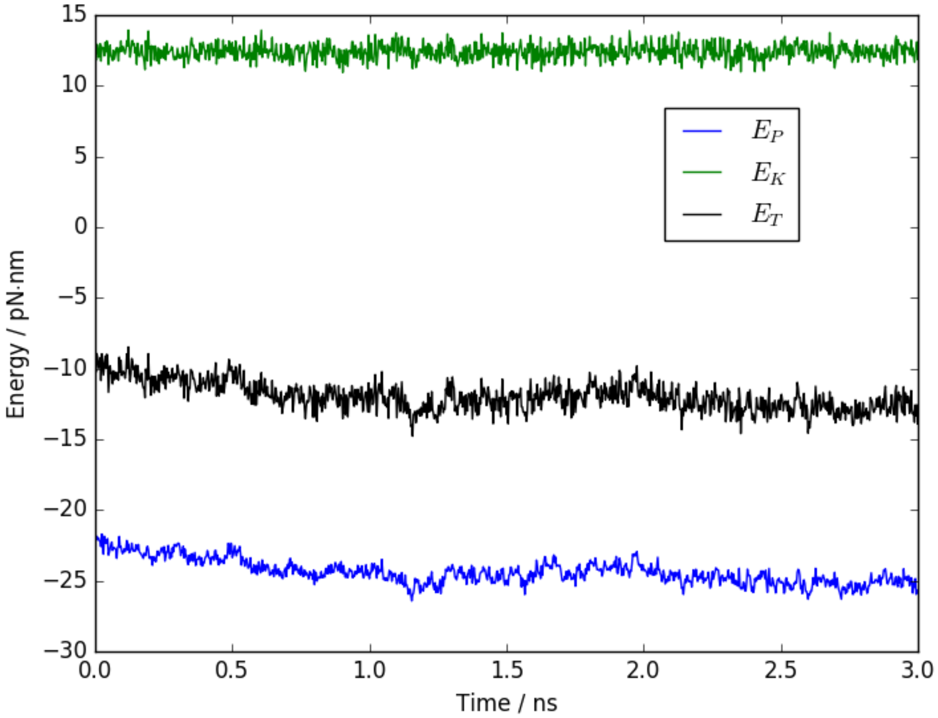


Figure 5.1: Total, potential, and kinetic energy over 3 ns for poly-d(A)·poly-d(T) stretched by 21.3 nm simulated in oxDNA

Initially, the minimisation procedure was tested. A stretching simulation was undertaken with 1000 steps of steepest descent minimisation after every stretching event. The

DNA underwent 100 stretching event-minimisation-simulation cycles and the final conformation was simulated for 3 ns and the potential and kinetic energies plotted. The final conformation was used as any build-up of unrelaxed energy over the stretching cycles would be visible.

A plot of the energy components over the 3 ns can be seen in Figure 5.1. The energy components remain approximately flat throughout the simulation, which indicates two things. First, the minimisation procedure was sufficient to relax the stretched DNA fragment, and the timestep is small enough not to drive more energy into the system through bad interactions with the immobilised terminal base pairs. Though there is some reduction in potential energy, it is small enough and decays quickly enough that there is no reason to believe that the qualitative structures obtained are untrustworthy – in other words, over time the system would not redistribute the stretching perturbation as it is at a constant energy.

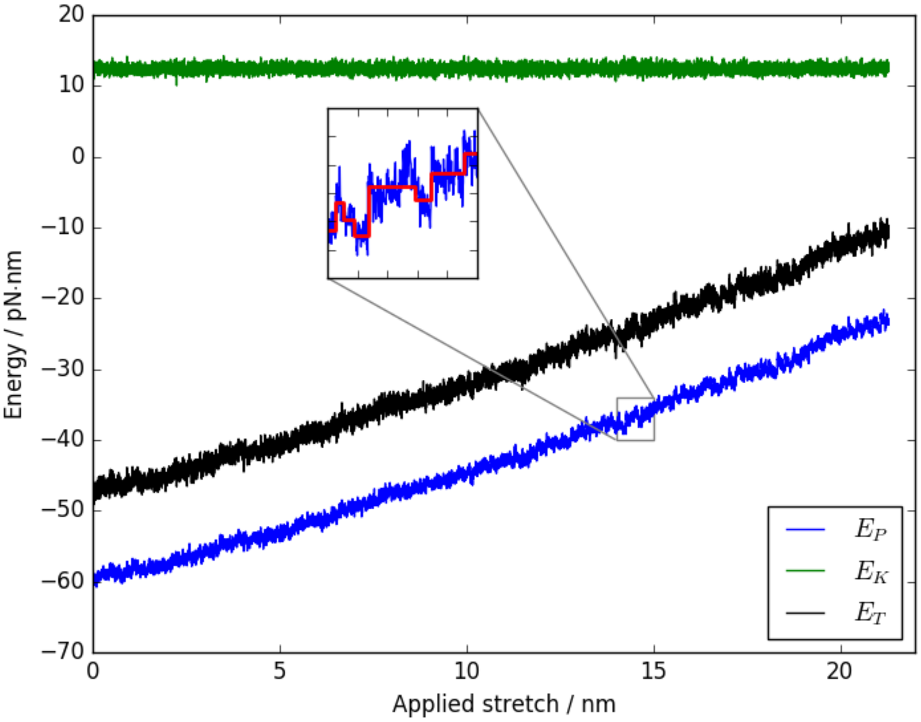


Figure 5.2: Total, potential, and kinetic energies for poly-d(A)·poly-d(T) stretched using oxDNA

With the timestep and minimisation verified, the simulation could be analysed in earnest. Again, the total, potential, and kinetic energies were plotted throughout the simulation, and the results are given in Figure 5.2.

The kinetic energy remains flat throughout the simulation, indicating that the thermostat was correctly moderating the temperature. However, the potential energy increases as the DNA is stretched. This is the expected behaviour for such a system. As the terminal bases are moved further away, the covalent bonds in the backbone are stretched, as are the non-bonded van der Waals interactions which are responsible for the stacking between base pairs. The strength of these is dependent on the separation of the particles and as they weaken the potential the particles experience weakens. As a result, the magnitude of the interaction decreases, corresponding to an increase in the negative potential energy.

Although on this scale the increase appears smooth, they are in fact stepwise increases in potential energy corresponding to the stretching events and bonds being broken and made, as would be expected: the contour length experiences a step change and therefore the potential energy experiences stepwise change also. The inset stepwise model was fitted using a hidden Markov model first used for determining trajectories in FRET experiments [163], and shows evidence of individual states in the simulation.

The overall system parameters of energy and temperature therefore behaved as expected, validating the simulation scheme.

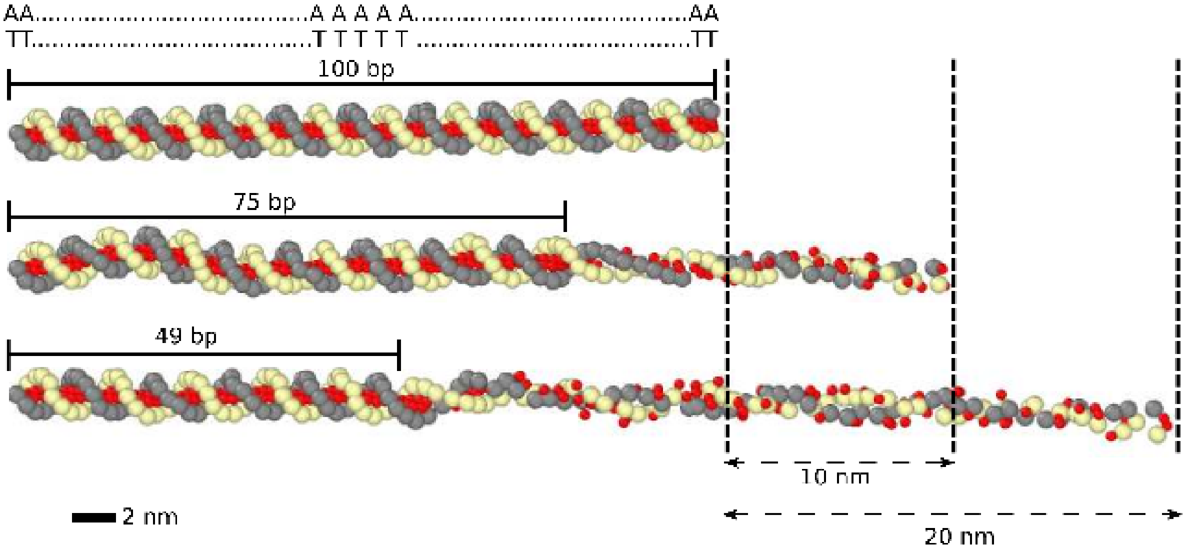


Figure 5.3: Simulated conformations of poly-d(A)·poly-d(T) simulated in oxDNA, stretched by a) 0 nm; b) 10 nm; c) 20 nm. Large grey and white spheres are backbone atoms while the base particles are shown in red.

Finally, the conformations of the DNA were examined at three times – the starting point of the simulation, halfway through the simulation, and at the end. The results of this can be seen in Figure 5.3.

These conformations are quite striking. Rather than being distributed along the length of the DNA, the stretching is localised to the end of the DNA that the stretching is happening from – instead of stretching along the full length like an elastic band, the DNA appears to be unwound from the moving end first. Adjacent to this is a short stretch of less perturbed DNA which acts as the interface between the highly disrupted region and the unchanged Watson-Crick double helix far from the stretching events.

Alongside this stretching, the base pairs themselves can be seen to break down. Initially the red dots of the bases are inside the double helix as is expected for dsDNA, but upon stretching they begin to turn outwards and break their hydrogen bonds, as is expected for DNA stressed in this way. However, the localisation of this is extremely suggestive that in experiments we may expect more disruption closer to the source of the stretching. It also suggests that it is possible to evaluate the localised perturbation that may be expected by looking at short stretches of DNA using all-atom molecular dynamics.

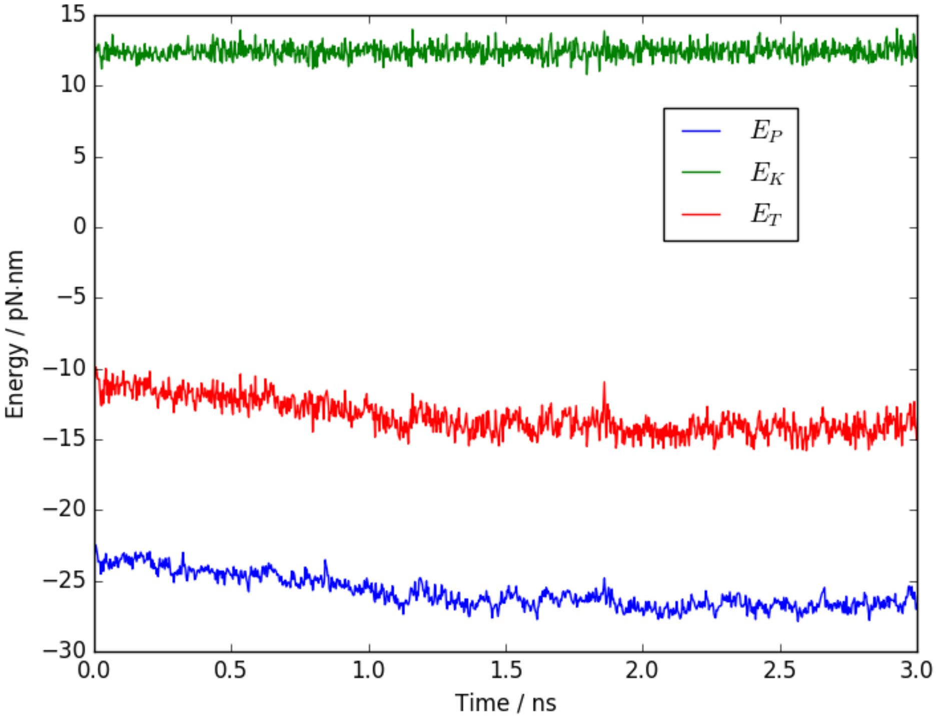


Figure 5.4: Total, kinetic, and potential energy for poly-d(C)·poly-d(G) over 3 ns when stretched by 20 nm in oxDNA

### 5.2.2 poly-d(C)·poly-d(G)

Once again the initial concern was to ensure that the simulation parameters were suitable for simulating a system constrained and perturbed by stretching. As before, the final stretching configuration was simulated for 3 ns and the energies plotted. This can be seen in Figure 5.4.

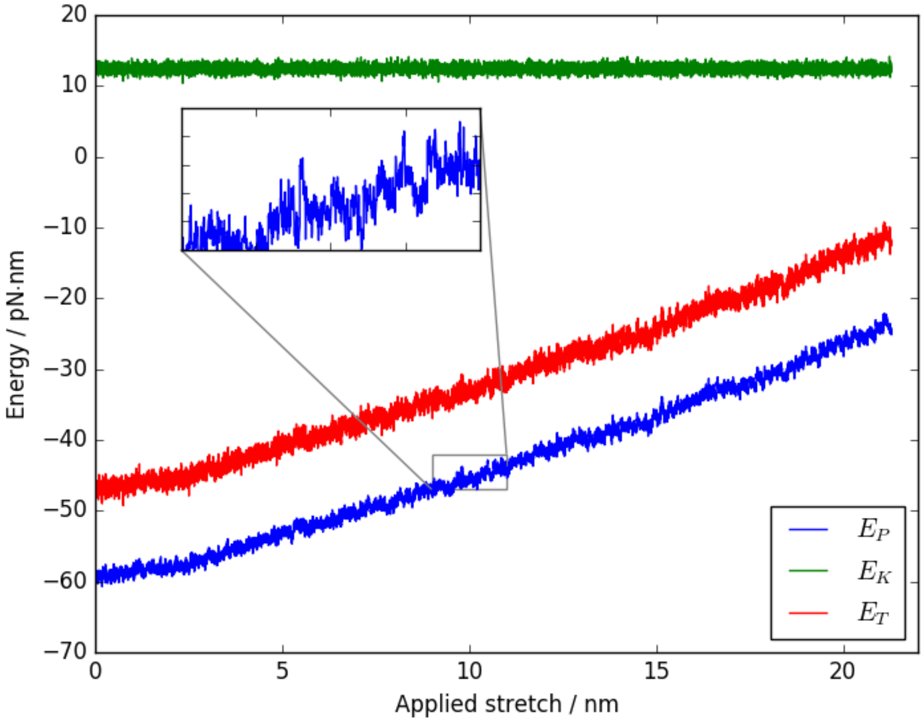


Figure 5.5: Total, potential, and kinetic energies over a full stretching simulation of poly-d(C)·poly-d(G) in oxDNA

Although in this graph the potential and therefore total energy decrease slightly through time, it is a small change relative to the energy of the system, and by the end of the 3 ns the energy has once again stopped changing. This indicates that while the absolute minimum energy conformation was not found during the minimisation procedure, the structure that was found was very close. Since the oxDNA simulations are used as qualitative, coarse-grained indicators of behaviour over the full contour length of the DNA, a small change at the start of a simulation is not of concern. No major structural changes are hinted at in this energy graph, and so the distribution of perturbation over a 300 ps simulation will be the same as over 3 ns for the purposes of this simulation. As for the poly-d(A)·poly-d(T) simulation, the lack of dramatic energy change also indicates that the minimisation and simulation procedure overall has been successful in relaxing excess energy out of the system.

The energies through time over the full stretching simulation were then plotted in Figure 5.5. Once again the potential energy increases with the applied stretch, while the thermostat keeps the kinetic energy constant. The energy increases are again stepwise. No energy spikes are seen, indicating as before that energy is not being incorrectly introduced to the system by interactions with the restraints on the terminal base pairs.

Finally, the structures themselves were qualitatively imaged as in Figure 5.6. Once again, the stretching leads to a collapse in the backbone of the DNA, with the bases themselves turning outwards and becoming exposed. As before, the stretching perturbs the DNA in three ways: closest to the moving base pairs the DNA is extremely perturbed and base pairs flip outwards, then there is a short region of minor disruption before the remaining section of unperturbed DNA. This correspondence with the previous simulation gives confidence that the conclusions drawn there were not due to base-pair sequence but are a legitimate mechanical property of the DNA itself.

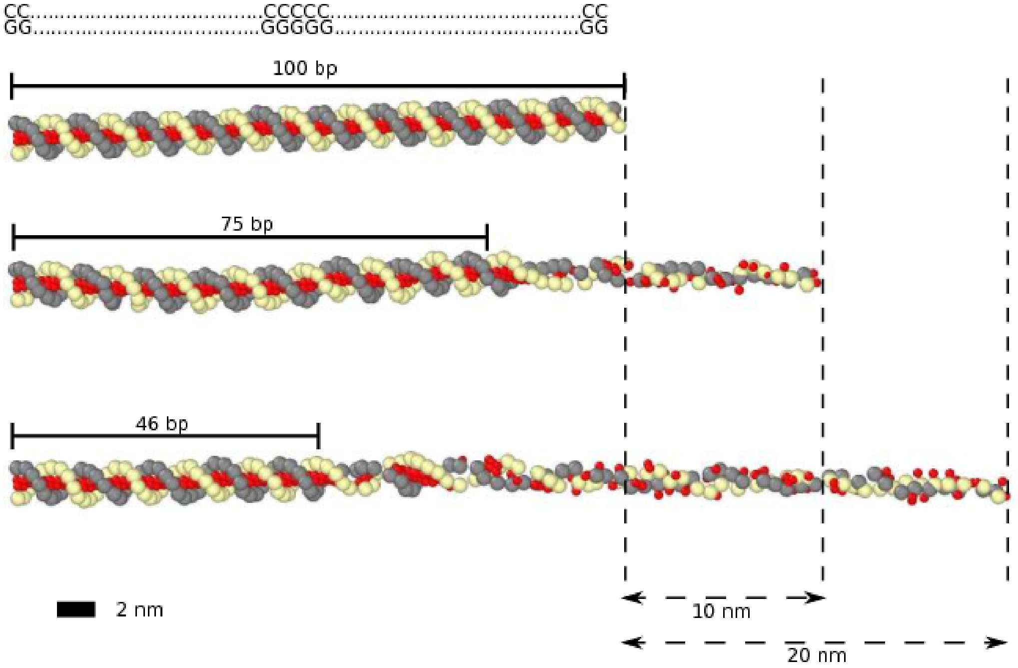


Figure 5.6: poly-d(C)·poly-d(G) stretched by a) 0 nm; b) 10 nm; c)20 nm in an oxDNA simulation

This localised stretching process is therefore predicted to be independent of sequence, or at least independent of the proportion of the sequence which is AT or CG. AT or CG content is a key predictor of many DNA parameters including melting temperature [164] and so it is striking that these simulations do not show an appreciable qualitative difference between the two over the length scale of 100 bp. In order to study more closely the sequence dependence of torque and sequence on DNA stretching behaviour, atomistic simulations are necessary and can justifiably be performed over short sequences to investigate the localised stretching regions found in the coarse-grained regime.

## 5.3 Amber Simulations

### 5.3.1 Minimisation

As described previously, extensive minimisation was carried out on the perturbed structures. This is more important to get right than in the oxDNA simulations: the number of bonds in an Amber structure is orders of magnitude larger than in oxDNA and so significant relaxation is required to be sure that none of them experience any bad contacts.

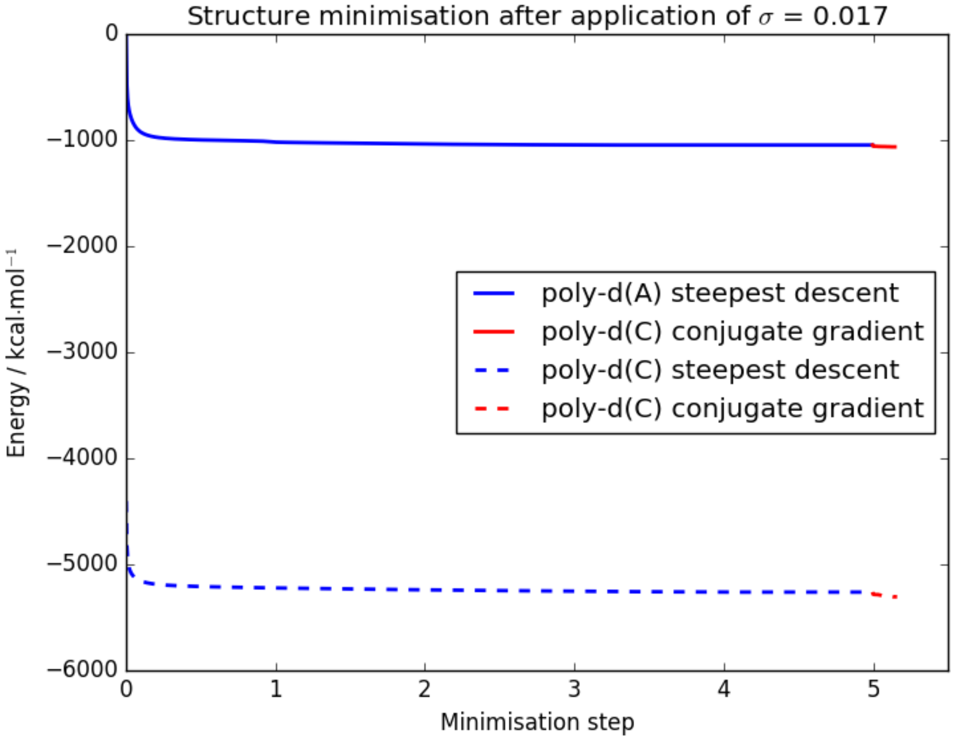


Figure 5.7: 0K minimisation of 24 bp dsDNA in Amber after application of a twist to one terminal base pair ($\sigma = 0.017$) relative to the initial energy of the poly-d(A)·poly-d(T) fragment

The initial perturbation in this simulation scheme is the rotation of the terminal base pair to generate torque on the structure, and the perturbation induced by this should be relaxed prior to any stretching or simulation. A long minimisation of at least $5 \times 10^6$ steps was undertaken, consisting of exactly $5 \times 10^6$ steepest descent steps followed by up to $10^6$ conjugate-gradient. In practice, the conjugate gradient algorithm terminated quickly after finding a minimum and thus far fewer than $10^6$ were performed. This long minimisation resulted in an energy profile seen in Figure 5.7.

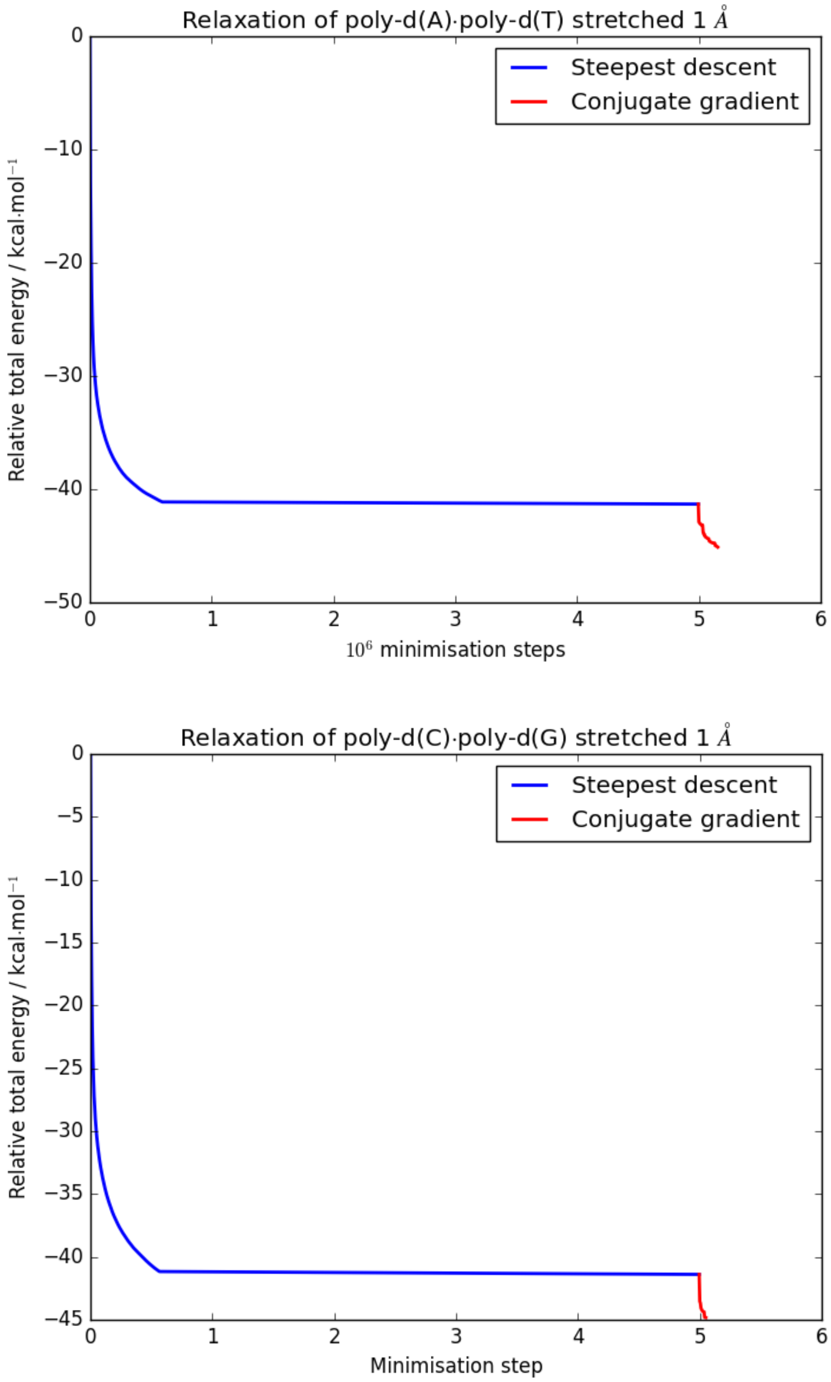
For both structures the energy quickly drops and its profile becomes flat as the steepest descent minimisation has reached convergence. At $5 \times 10^6$ steps the conjugate gradient algorithm takes over and reduces the total energy still further before quickly reaching its limit. The final decrease in energy is small: 17.43 kcal·mol$^{-1}$ for poly-d(A)·poly-d(T) and 44.08 kcal·mol$^{-1}$ poly-d(C)·poly-d(G), corresponding to 0.75 eV per molecule and 1.90 eV per molecule respectively. In quantum chemistry structure minimisation, a good level of energy convergence is often taken to be 1 meV per atom, and with over a thousand atoms per structure the minimisation performed here is at that level. The energy minimisation for the rotation of the terminal base pair is therefore very strict.

An additional feature of Figure 5.7 is the relative scale of the energies of the two structures. The GC-rich structure has a final relaxed energy some 5000 kcal·mol$^{-1}$ below that of the AT-rich sequence. This is in accordance with what would be expected. The hydrogen bonds in a GC pair are both more numerous (3 compared to 2) than those in the AT base pair, and they are stronger. In addition it is well known that the stacking interactions for GC rich DNA are stronger than those in an AT rich segment, leading to the different melting temperatures in [164]. Thus the potential energy of the GC rich fragment should be greater as the DNA sits in a deeper potential well.

The second perturbation experienced by the DNA is the stretching event, where one base pair is moved 1 Å away from the other end of the duplex. This again needs careful minimisation – too little and the resulting forces will be large enough to give unphysical results. As the previous scheme of $5 \times 10^6$ steepest descent steps followed by conjugate gradient minimisation worked well for the rotation, it was tried for the stretching events. Once again, the total energy of the molecule was plotted for two sequences – poly-d(A)·poly-d(T) and poly-d(C)·poly-d(G) – as they underwent the minimisation.

In Figure 5.8 is shown the energy profile for the structures being minimised after being stretched 1 Å. Both see sharp reductions in total energy during the steepest descent phase followed by a further reduction due to the superior conjugate-gradient method. As in the rotation minimisation, the final drop in energy is very small – 3.8

and 3.4 kcal·mol$^{-1}$ for the poly-d(A)·poly-d(T) and poly-d(C)·poly-d(G) respectively, well below 1 meV/atom and therefore extremely strict.



a)

b)

Figure 5.8: Energy during a 0K minimisation for a) poly-d(A)·poly-d(T) and b) poly-d(C)·poly-d(G) stretched by 1 Å in Amber. Energies are relative to each structure's initial energy

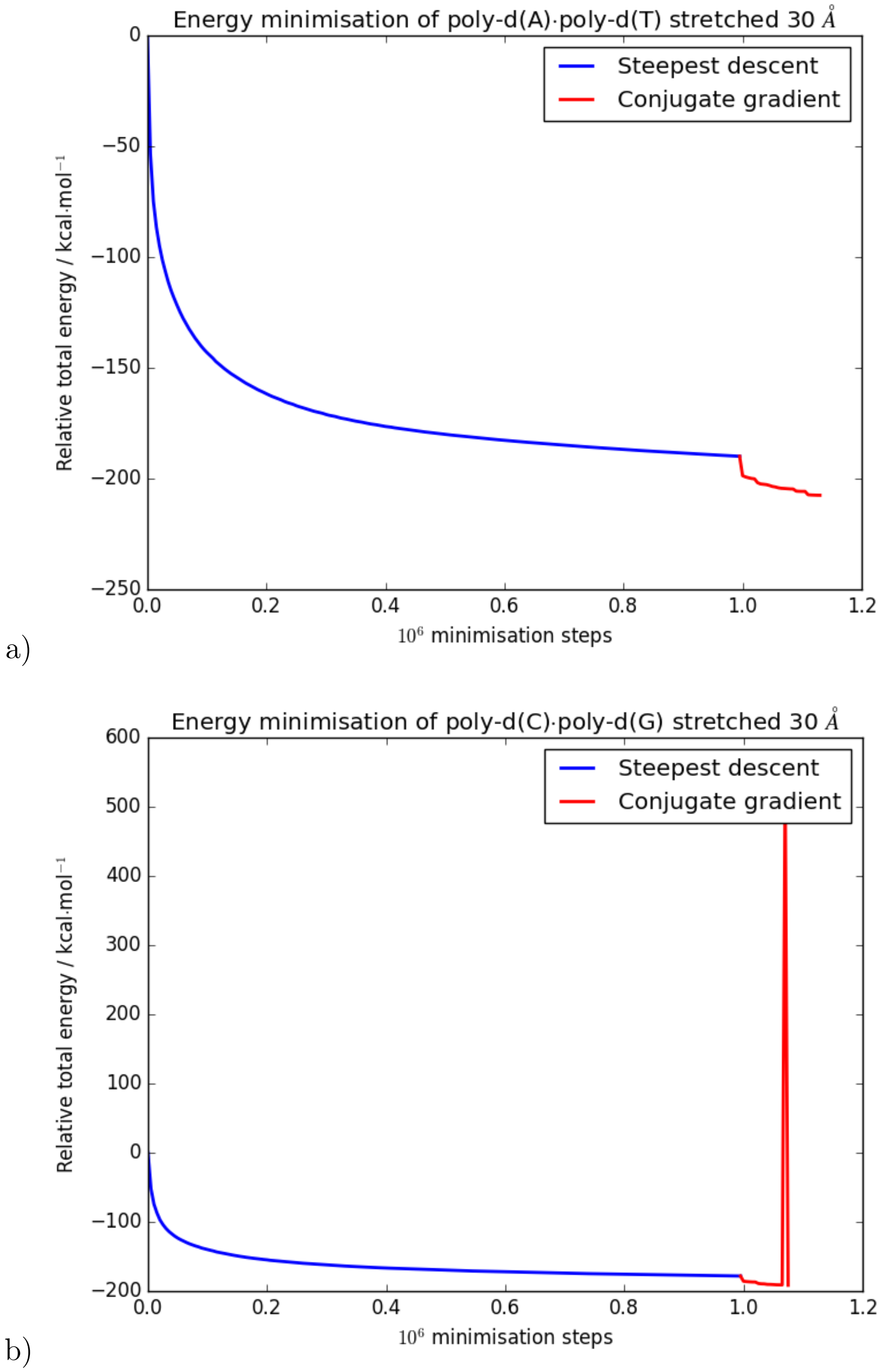


Figure 5.9: Minimisation with $10^6$ steepest descent steps followed by conjugate-gradient for a) poly-d(A)·poly-d(T) and b) poly-d(C)·poly-d(G) stretched in Amber by 30 Å. Energies are relative to the structures' starting energy

However, this minimisation routine took considerable computational expense and with so many separate simulations to run it was significantly lengthening the run time.

Instead of $5 \times 10^6$ steepest descent steps $10^6$ were tried, again followed by conjugate-gradient minimisation. To test this thoroughly, two structures stretched by 30 Å were minimised in this way.

The results are seen in Figure 5.9. The precipitous energy relaxation during steepest descent is visible as is the subsequent change when conjugate-gradient is employed. Once again the change due to conjugate-gradient after steepest descent is low – 27.7 kcal·mol$^{-1}$ in Figure 5.9a and 21.8 kcal·mol$^{-1}$ in Figure 5.9b. This again works out to be below 1 meV/atom. This is important as it demonstrates that the structures are already well minimised for this application prior to the conjugate-gradient scheme being introduced and very little energy remains to be relaxed out when moving to the different minimisation routine.

One striking feature of 5.9b is the dramatic spike in energy at the end of the conjugate-gradient minimisation steps. This is however not a chief concern. The energy returns the next step to the level it was before indicating that this jump is a result of numerical instability. The likelihood is that the system was so well minimised that when calculating the conjugate-gradient and the step size, the algorithm broke down and gave an incorrect value. However, at the next cycle the error was put right immediately and the algorithm terminated as it could find no lower energy configuration.

#### 5.3.1.1 Equilibration

A full simulation was performed and the energies and temperatures analysed to assess the need for equilibration molecular dynamics too be performed. Similarly to the minimisation, the structures at 1 Å and 30 Å stretches were investigated.

The results for the 1 Å simulation are shown in Figure 5.10. The equilibration phase for all properties investigated are extremely brief and represents only the first frame of the simulation. Thereafter, the potential and kinetic energies remain close to a constant average vale. This indicates that the structure is well equilibrated from the start and at this stage a dedicated equilibration step is not necessary.

The results for the structure stretched by 30 Å is presented in Figure . Once again the values almost instantly converge to an average value which is then maintained until the end of the simulation. Once again in this case it is unnecessary to perform a specific equilibration simulation. The first frame, which is the only one which is affected by the equilibration, does not need to be excised from the simulation prior to analysis as its effect on the average structure and values is minimal.

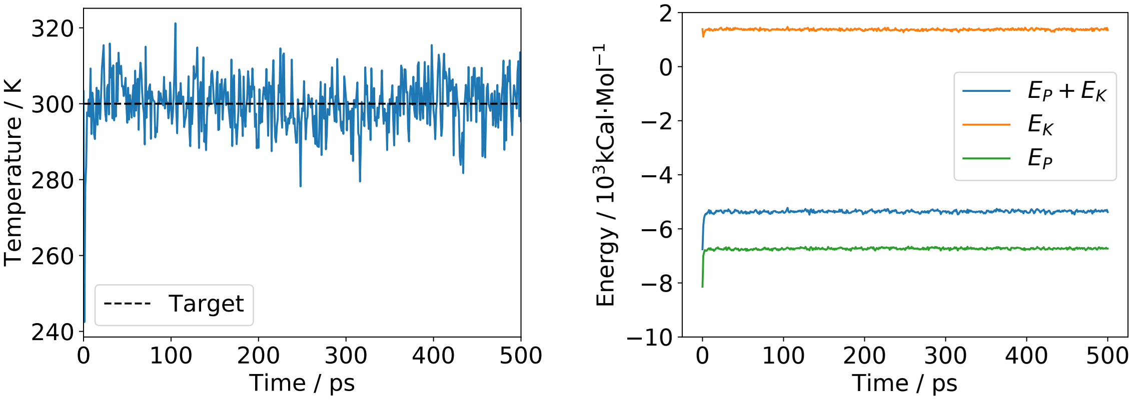


Figure 5.10: Temperature and energy over a 500 ps simulation of poly-d(A)·poly-d(T) stretched by 1 Å. After a very brief initial equilibration period, the temperature, kinetic energy, and potential energy all stay close to an average value for the remainder of the simulation. In this case therefore an equilibration MD stage is not needed, and as the initial equilibration is so brief it is not necessary either to discard the first molecular dynamics frame before analysis.

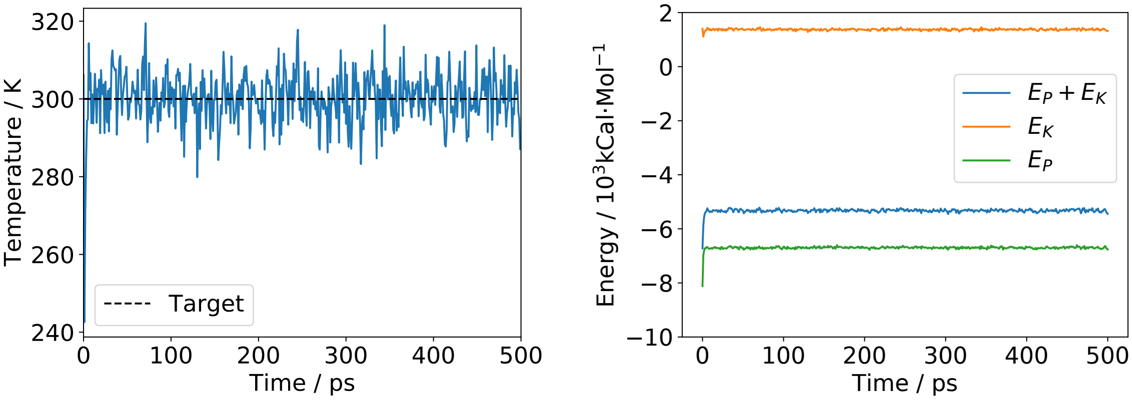


Figure 5.11: Temperature and energy over a 500 ps simulation of poly-d(A)·poly-d(T) stretched by 30 Å. As before the investigated parameters almost instantly reach their average values, and there is no evidence of the structure seeking a lower energy state as the potential energy trace is flat. Once again no dedicated equilibration step is needed, and the first frame can safely remain as part of the analysis as its contribution is small.

As the structures were found to have reached equilibrium in both of the extreme cases, no equilibration step was performed, and no frames were removed from the simulations prior to analysis.

### 5.3.2 poly-d(A)·poly-d(T)

The production molecular dynamics was then performed on the four structures introduced previously. For poly-d(A)·poly-d(T), the initial structure can be seen in Figure 5.12, and it can be seen that consistent with oxDNA simulations and previous work [49] that the bases have begun to break and the backbone to collapse. This base pair breaking is a key feature of the stretching pathway of DNA and to quantify this the hydrogen bonds which hold the base pair together were analysed throughout the simulation.

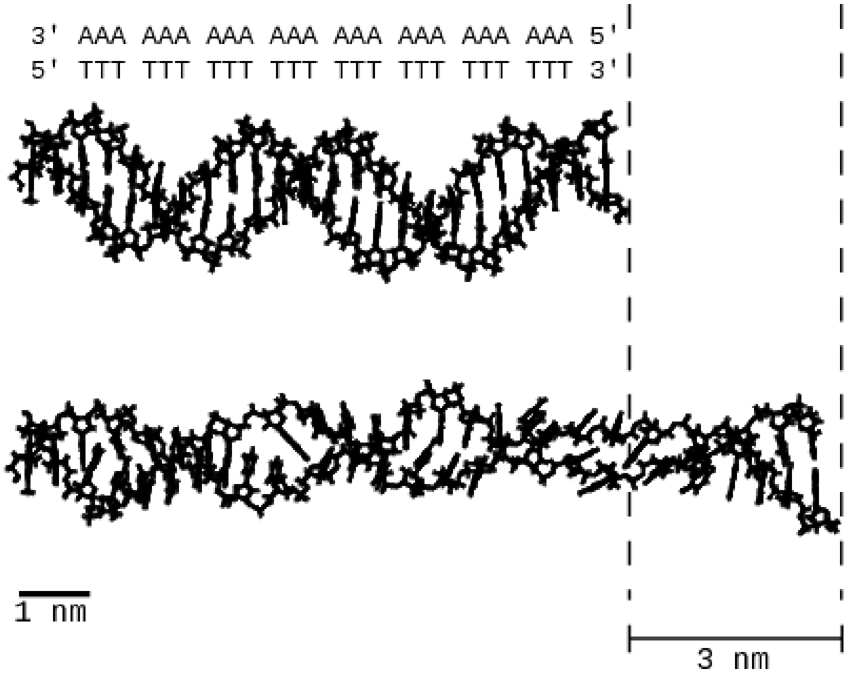


Figure 5.12: Structures from the stretching simulation of untwisted poly-d(A)·poly-d(T). Upper: the starting structure; lower: the structure stretched by 30 Å.

Using the code provided by Dr Noy and a bespoke script written by the author, the canonical and non-canonical hydrogen bonds were plotted as a heatmap through time on separate axes. In both cases, `cpptraj` [165] was used to create a matrix of hydrogen bonds, wherein a 0 indicates that no hydrogen bond exists between the two atoms and a 1 exists that a hydrogen bond is present in that frame. In Dr Noy's software, the hydrogen bonds between canonical Watson-Crick base pairs only are selected – hydrogen bonds between bases that would not normally form them are ignored. In the author's own analysis procedure, the situation is reversed – canonical interactions are excluded. Two plots for each separate frame of a simulation are thus produced –

one which charts the Watson-Crick base pairing and one which indicates where new, unexpected hydrogen bonds have been found.

In each case, the information is plotted on a base pair level and given as an average number of bonds present. Thus in the canonical scheme, an AT base pair having a value of 2 canonical hydrogen bonds and 0 non-canonical hydrogen bonds indicates that that base pair is interacting via hydrogen bonds in a completely standard way. If the values are 1 canonical and 1 non-canonical hydrogen bond, the base pair retains one of its Watson-Crick hydrogen bonds but one of the bases has formed a new hydrogen bond in addition. However it is not plotted here where that hydrogen bond has been formed.

Dr Noy's analysis code averages over 10 frames instead of processing each snapshot as an individual. This gives more information about the occupancy of the bonds. For example, a value of 1.5 may indicate that on average one and a half bonds are present – this could be one bond always present and the other present half the time, or another more complex set of interactions.

In the author's analysis, the hydrogen bonds are summed for each base pair and halved to avoid double-counting bonds. A single hydrogen bond between base A and B increases the reported hydrogen bonds by one half for each of A and B's base pairs. As well as ensuring that the sum of the hydrogen bonds is the number of non-canonical hydrogen bonds present, this serves also to compress the range of the non-canonical data set, making visualisation more straightforward and limiting the number of colours needed to represent the heat map.

The result of this analysis for the poly-d(A)·poly-d(T) is seen in Figure 5.13. Due to the size of the dataset, only the extremes of applied $\sigma$ are shown.

In Figure 5.13 it can be seen that for each $\sigma$ the stretching induces breaking of the canonical Watson-Crick base pair hydrogen bonds, as expected. The bond breaking appears weakly dependent on the applied torsion, with the canonical bonds breaking at different extensions but all in the range 5-10 Å. Alongside the bond breaking, occurring at the same extensions and locations as the breakdown of Watson-Crick bonds, is a rise in the number of non-canonical hydrogen bonds between the base pairs. As these are made as the canonical bonds break, it seems likely that it is the same hydrogen bonding sites used for both interactions, and that the bases are stabilising the helix by forming new hydrogen bonds as their preferential ones break.

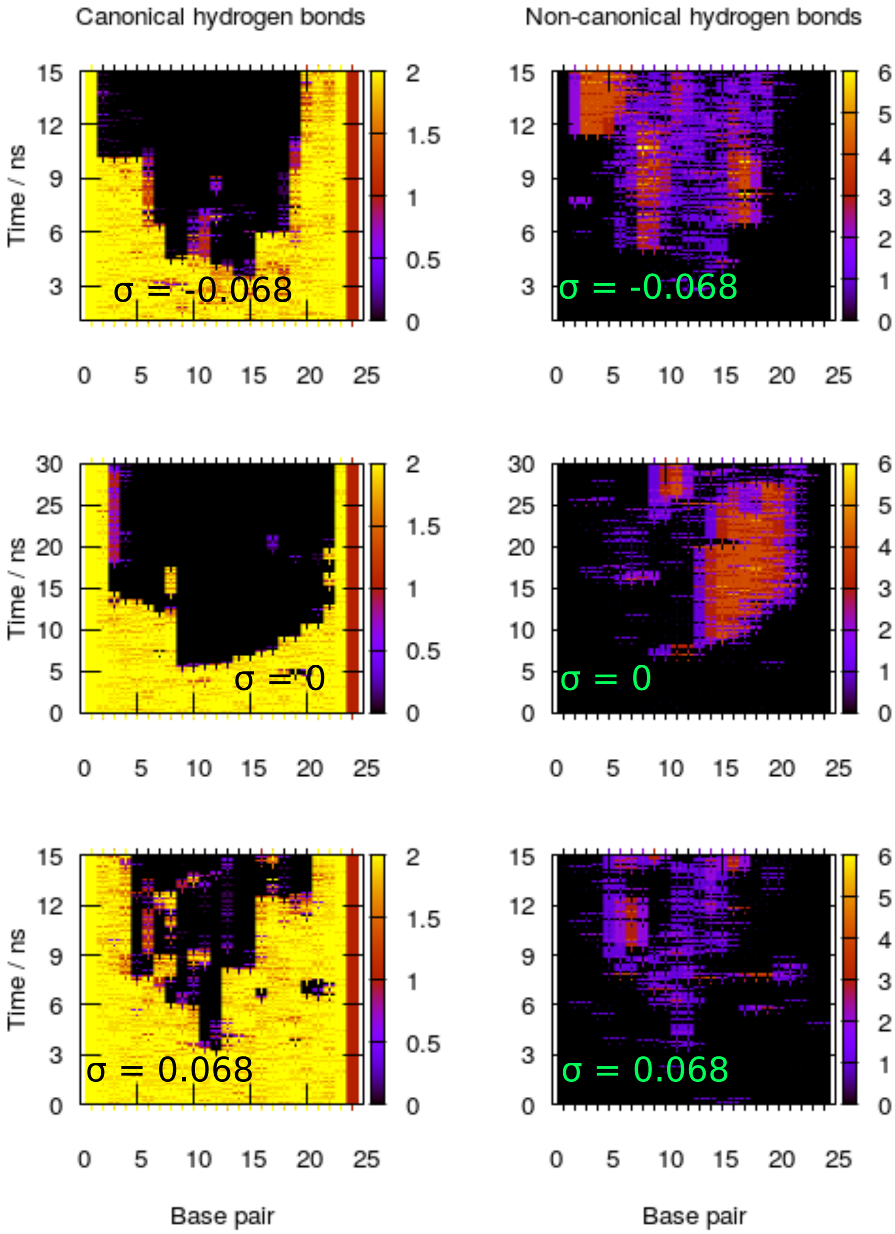


Figure 5.13: Hydrogen bonds for stretched and twisted poly-d(A)·poly-d(T) in Amber with $\sigma = 0, \pm 0.068$. Each $x$ axis represents base pair number, and the $y$ axis is the simulated time. A sudden stretch of 1 Å is applied every 500 ps for the $\sigma = \pm 0.068$ simulations and every 1 ns for the $\sigma = 0$ simulation.

An eyecatching localised patch of higher intensity can be seen in the non-canonical hydrogen bonds for the $\sigma = 0$ simulation. This persists for considerable stretching duration, suggesting that a semi-stable structure has been formed but is then torn apart by increased stretching. This is borne out by the reduction in non-canonical hydrogen bonds at the end of the simulation: these exotic interactions were no longer possible at the end of the simulation.

The correspondence between the non-canonical and canonical hydrogen bond heatmaps is a feature of each simulation. In the $\sigma = 0.068$ simulation, a temporary loss of canonical hydrogen bonds at base pair 8 around 20 Å stretch is mirrored exactly by an uptick of non-canonical hydrogen bonds in the same location at the same stretch. As the canonical bonds reassert themselves later, the non-canonical bonds are broken once more.

Another striking feature of these graphs is that the disruption to base pairing tends to start in a particular place and spread outwards from there. This localised bond breaking and making was suggested at by the oxDNA simulations, but here it is on a base pair level. This is again somewhat expected from previous work – other simulations have shown that under extreme conditions the DNA tends to retain as much canonical B-DNA as possible [53, 54] – but the formation of new local structures in rotationally clamped and stretched DNA has not been reported in the literature.

It is possible, as can be seen in Figure 5.13, for a base pair to have both canonical and non-canonical interactions. One way to visualise this is to colour a DNA structure according to its interactions. In order to do this, a single structure was needed to represent the simulation. One way of finding such a structure is to generate an average – that is, literally to average the co-ordinates visited by each atom and hence build a structure. This is in general a poor way to visualise highly mobile systems – for example, averaging a protein folding would give physically meaningless results, and averaging a highly flexible residue may lead to atoms overlapping. For a short DNA construct which is constrained at each end, however, the limited available states mean that an average is likely to be a reasonable reflection of the DNA conformation, and while novel analysis performed on such an average may not be reliable, the generated average co-ordinate set is a perfectly good candidate for colour coding to visualise bond distribution as performed here.

Thus, the colour coding was done for the average structure of poly-d(A)·poly-d(T) stretched by 25 Å using both 1- and 2-D heatmaps coded in the following ways:

| 0-1 | 1-2 | 2+ |
|---|---|---|
| Grey | Yellow | Green |

Table 5.1: Heatmap coding for the canonical bond structure heatmaps

| 0-1 | 1-3 | 3+ |
|---|---|---|
| Grey | Red | Blue |

Table 5.2: Heatmap coding for the non-canonical heatmaps

| | 0-1 | 1-2 | 2+ |
|---|---|---|---|
| 0-1 | Grey | Yellow | Green |
| 1-2 | Orange | Pink | Lime |
| 2-3 | Red | Purple | Turquoise |
| 3-4 | Cyan | | |

Table 5.3: Heatmap coding for the full colour hydrogen bond heatmaps. Columns are canonical hydrogen bonds, rows are non-canonical hydrogen bonds

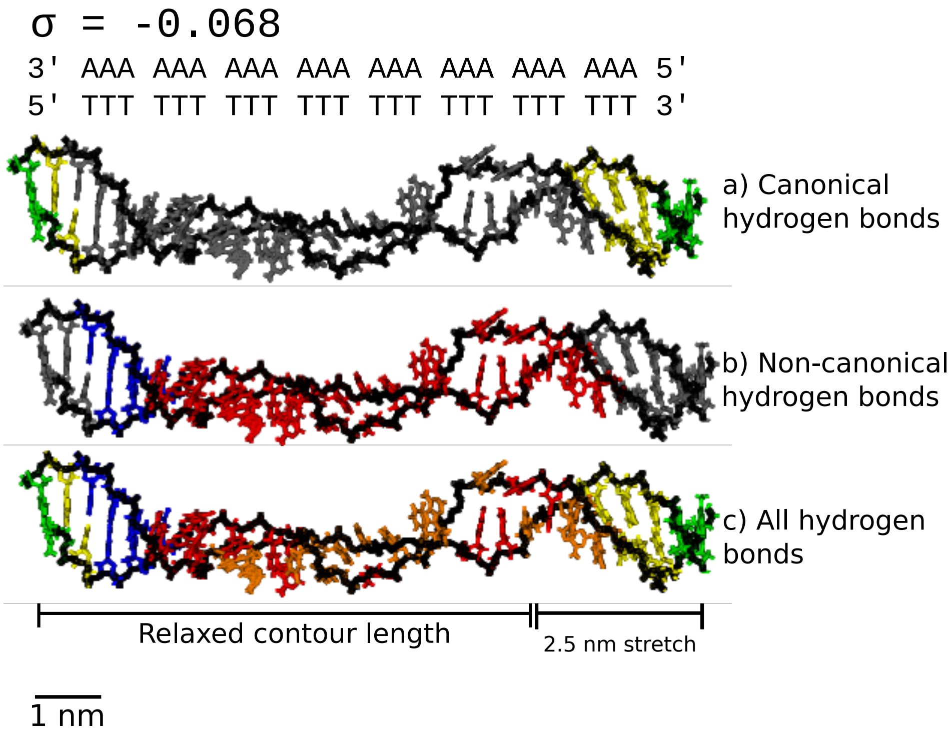


Figure 5.14: Heatmaps showing the hydrogen bonding of poly-d(A)·poly-d(T) with an applied stretch of 25 Å and $\sigma = -0.068$. From top to bottom: canonical interactions, non-canonical interactions, and full colour

The results for $\sigma = -0.068$ are shown in Figure 5.14, $\sigma = 0$ in Figure 5.15 and $\sigma = 0.068$ in Figure 5.16.

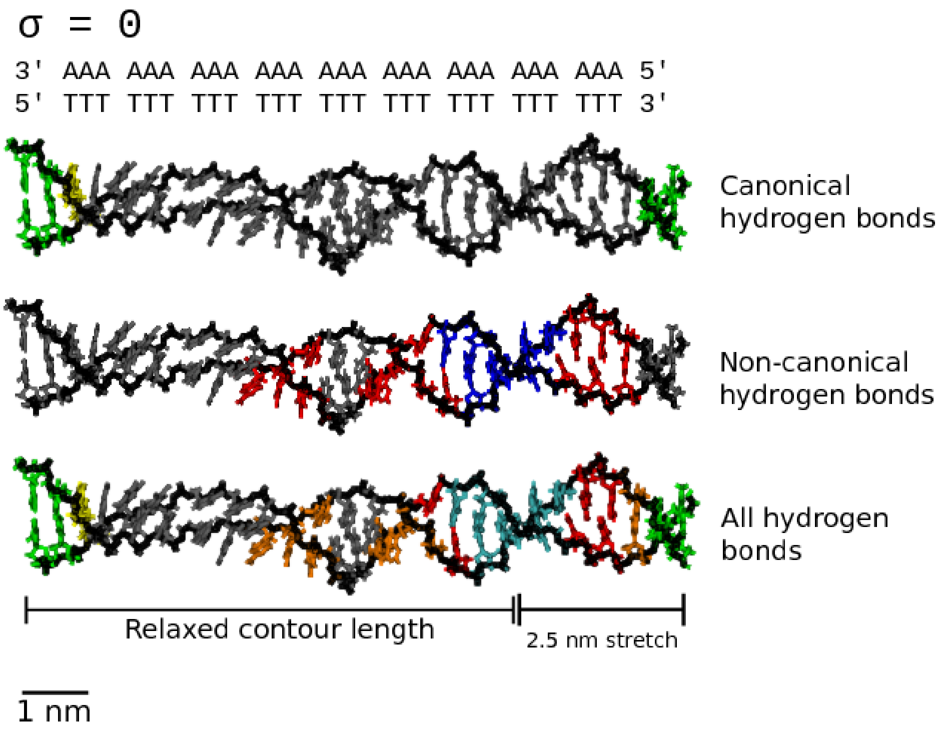


Figure 5.15: Structure heatmaps for hydrogen bonds for poly-d(A)·poly-d(T) stretched 25 Å and with $\sigma = 0$

As in Figure 5.13, it is clearly seen that the non-canonical hydrogen bonds are formed where the canonical hydrogen bonds break down. The backbone, shown in black, can be seen to have collapsed such that the bases are oriented away from the helical axis, though again this is not uniform across the structure. The clamped end base pairs appear also to keep their neighbours in a canonical form – without the ends being clamped it is likely that the DNA would by that stage have dissociated completely.

Finally, the stacking interactions were examined on a coarse basis through plotting of the total van der Waals energy. The van der Waals energy through time was plotted and is shown in Figure 5.17. The values shown in the graph increase as the stacking interactions break down, but are extremely noisy showing no real cohesion. Each of the structures shows a decrease in van der Waals energy at around 20 Å extension, but since this is a very coarse measure, no real information can be drawn from it. This strengthening of the stacking interactions does not correlate with any hydrogen bonding information, however it does suggest there may be new stacking interactions forming. Without detailed base pair van der Waals (VdW) information, however, going beyond saying that these drops in the VdW energy are suggestive would be pure speculation.

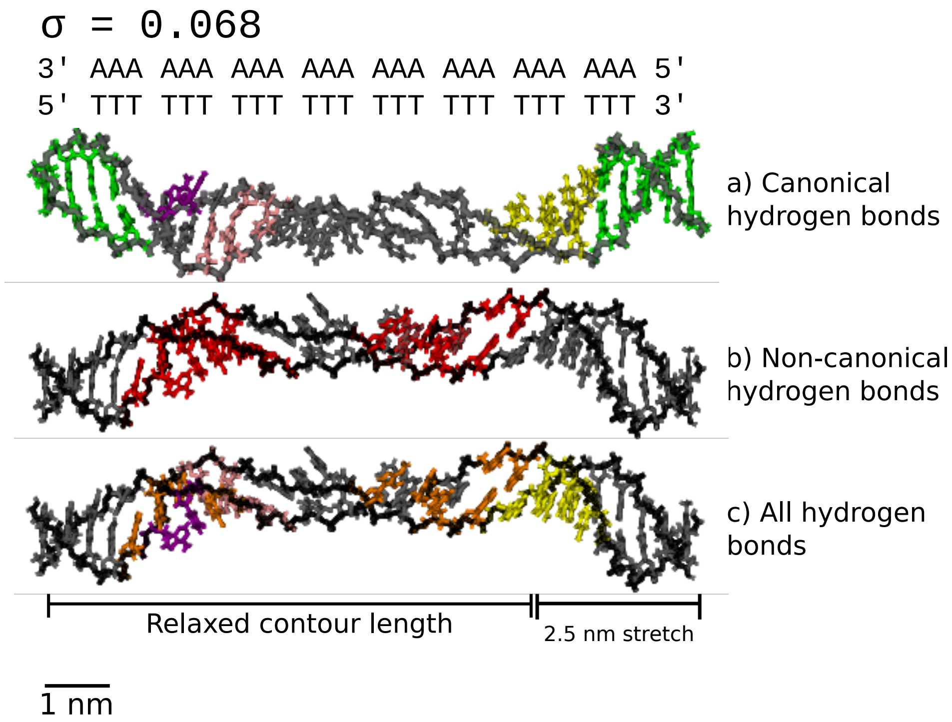


Figure 5.16: Structural heatmaps showing canonical hydrogen bonds, non-canonical hydrogen bonds, and all hydrogen bonds for poly-d(A)·poly-d(T) stretched by 25 Å with $\sigma = 0.068$

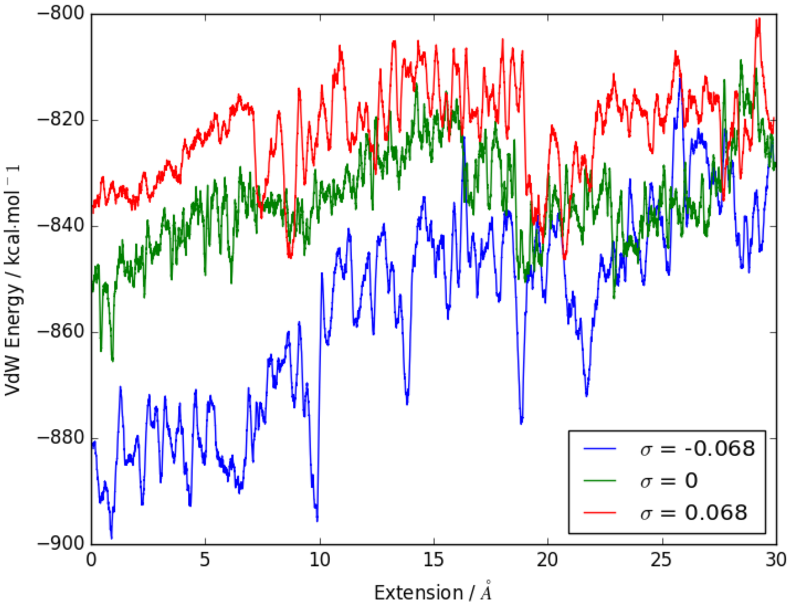


Figure 5.17: van der Waals energy through time for poly-d(A)·poly-d(T) over a simulation stretching it to 30 Å with $\sigma = 0, \pm 0.068$ with a 100 point rolling average applied

### 5.3.3 poly-d(AT)·poly-d(AT)

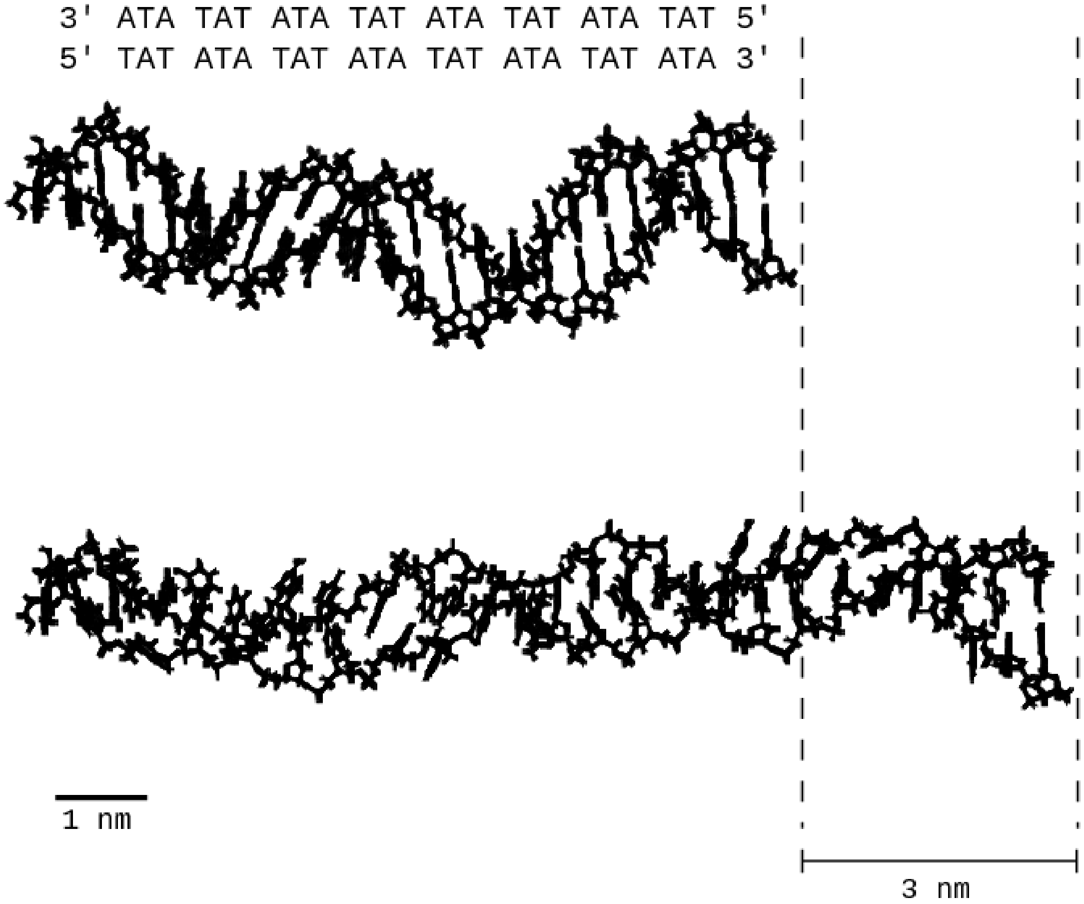


Figure 5.18: Upper: starting structure for the 15 ns poly-d(AT)·poly-d(AT) simulation with $\sigma = 0$. Lower: structure at a stretch of 30 Å

The same procedure was then followed for poly-d(AT)·poly-d(AT). Again, the system was built, minimised, and simulated. The initial structure and conformation found at a stretch of 30 Å can be seen in Figure 5.18. Once again, the backbone appears to be collapsing and leaving the bases to turn outwards, evidenced by the helix getting thinner. However, only so much information about a structure can be determined by eye.

Again, the extent to which the base pairs lost their canonical hydrogen bonds was assessed by way of a heatmap, and again the same was generated for the non-canonical hydrogen bonds. These heatmaps are presented in Figure 5.19.

Once more, the Watson-Crick canonical base pairs break down as the system is put under more torsion, and again the torsion applied appears to have some effect. Specifically, the overtwisted DNA retains its canonical interactions longer than either the neutral or undertwisted DNA. This is again in agreement with previous experimental work [96] which showed that stretching with rotational freedom led to the DNA overtwisting of its own accord. Thus a stretch after applied overtwisting would be expected to maintain the structure for longer. When the inevitable breakdown does occur, it is again matched by the appearance of non-canonical hydrogen bonds which serve to help stabilise the new structure.

The most striking example of these non-canonical hydrogen bonds, both for this sequence and this thesis as a whole, is the heatmap showing the non-canonical hydrogen bonds for undertwisted and stretched poly-d(AT)·poly-d(CG) in Figure 5.19. There, there is a rise of non-canonical hydrogen bonds as the Watson-Crick base pairing breaks down, and these non-canonical hydrogen bonds form a regular, periodic profile, as the intensity of non-canonical bonds reaches a peak every 5-6 base pairs, with a decline in between to a minimum after 3 base pairs and a steady increase to another maximum. This periodicity is highly suggestive of a novel structure being formed, especially so when the periodicity itself holds together and persists for the full length of the stretching simulation, appearing after roughly 15 Å of stretch and remaining up until the end of the simulation.

A similar thing happens in the untwisted simulation also, but these periodic maxima of non-canonical bonding do not stay spatially fixed and move in the earlier part of the simulation. At the end of the simulation, where the periodic peaks have asserted themselves and remained for the last 5 stretching events, the peaks are closer together and come around every 3-4 base pairs.

Taken together, these heatmaps are indicative of a novel structure created under tension in poly-d(AT)·poly-d(CG). In the 500 ps/Å/$\sigma$ simulations, DNA overwound to $\sigma = 0.068$ does not show this behaviour, while DNA underwound to $\sigma = -0.068$ exhibits a regular periodic structure asserting itself rapidly and stably over a long range of applied extension. In the replicas, however, the picture is more mixed, as periodic structures are seen for all $\sigma$, and the $\sigma = -0.068$ does not present the behaviour as strongly.

The poly-d(AT)·poly-d(AT) simulations are the only ones in which the replicas differ substantially from the longer simulation. One such replica is presented in Figure 5.20. The periodic structure remains in view, though it is not as twist dependent as in the main simulation, but is instead visible for all three $\sigma$ values shown. This suggests that the rate of pulling is also a factor in structures that can be formed and their stability. Intuitively this makes sense – a very quick rate of pulling may not allow the DNA to find stable structures. However, all the replicas show the same behaviour to a certain degree. It appears that for poly-d(AT)·poly-d(AT) such a structure is favourable upon torsionally constrained overstretching.

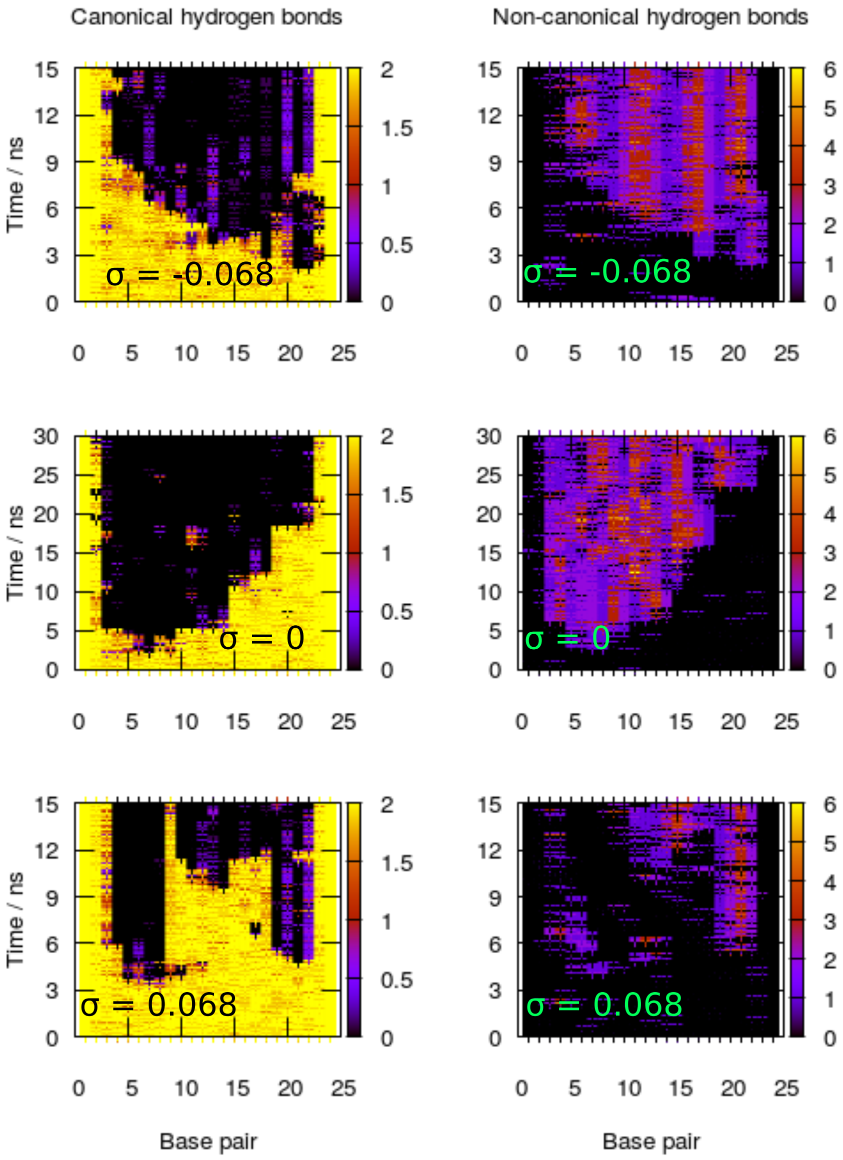


Figure 5.19: Hydrogen bond heatmaps of canonical and non-canonical hydrogen bonds for poly-d(AT)·poly-d(AT), with $\sigma = 0, \pm 0.068$. Once again the $y$ axis represents the simulated time while the horizontal axis shows the number of the base pair. An instantaneous stretch of 1 Å is applied every 500 ps for the $\sigma = \pm 0.068$ simulations and every 1 ns for the $\sigma = 0$ simulation.

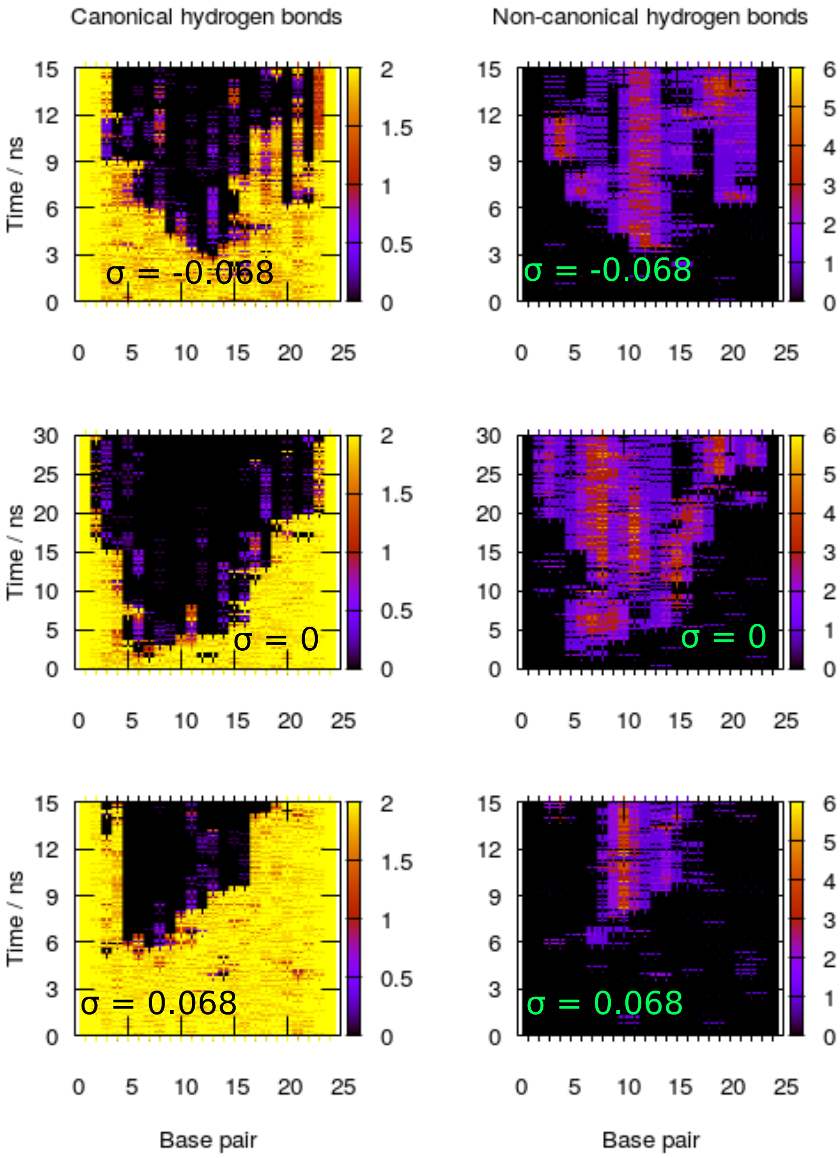


Figure 5.20: 100 ps/Å replica of poly-d(AT)·poly-d(AT) with $\sigma = 0, \pm 0.068$. Similar behaviour to that in Figure 5.19 is seen for all $\sigma$. The structures appear less stable than those in Figure 5.19.

These periodic structures can be visualised very clearly on the DNA structure itself. In Figures 5.21, 5.22, and 5.23 the structures are coloured as in Tables 5.1, 5.2, and 5.3. The periodic structures in both Figures 5.21 and 5.22 is very clearly visible in both the two-colour and full-colour plots, while little discernible pattern and areas of no hydrogen bonding at all in Figure 5.23 suggests that the DNA is behaving in a less structured way.

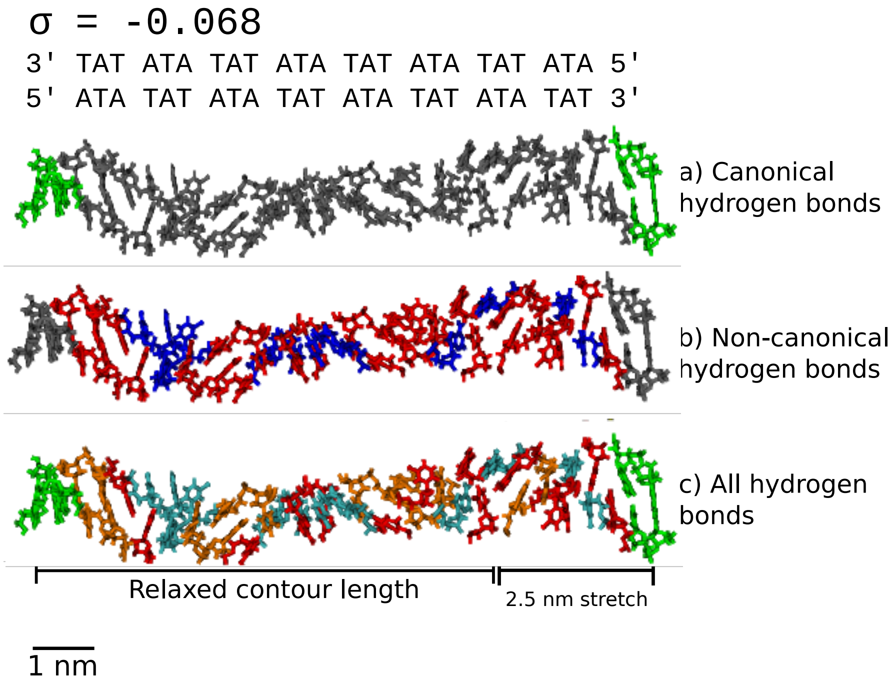


Figure 5.21: Canonical, non-canonical, and all hydrogen bonds for poly-d(AT)·poly-d(AT) with $\sigma = -0.068$ stretched 25 Å

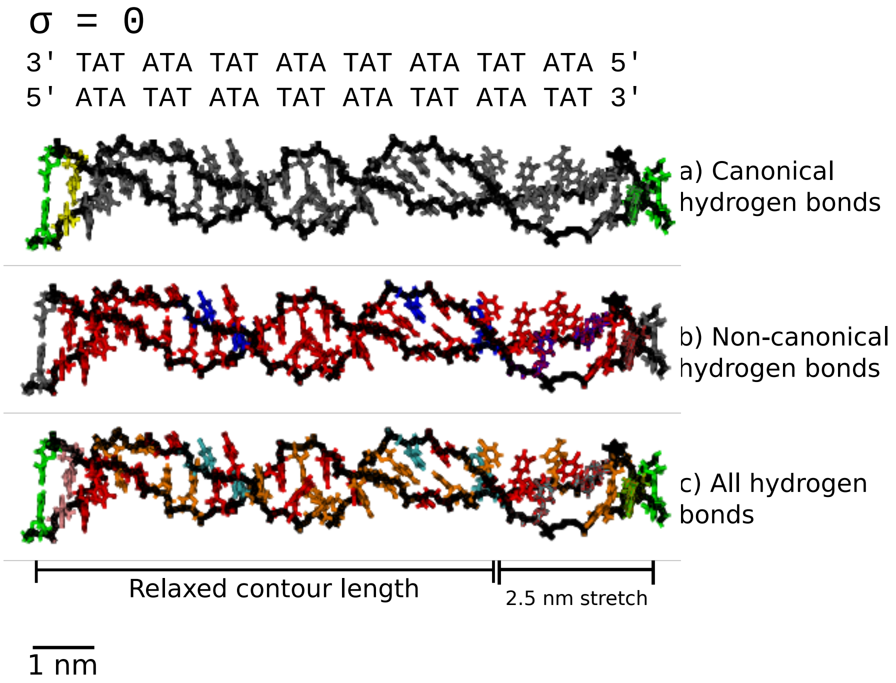


Figure 5.22: Canonical, non-canonical and all hydrogen bonds for untwisted poly-d(AT)·poly-d(AT) stretched by 25 Å

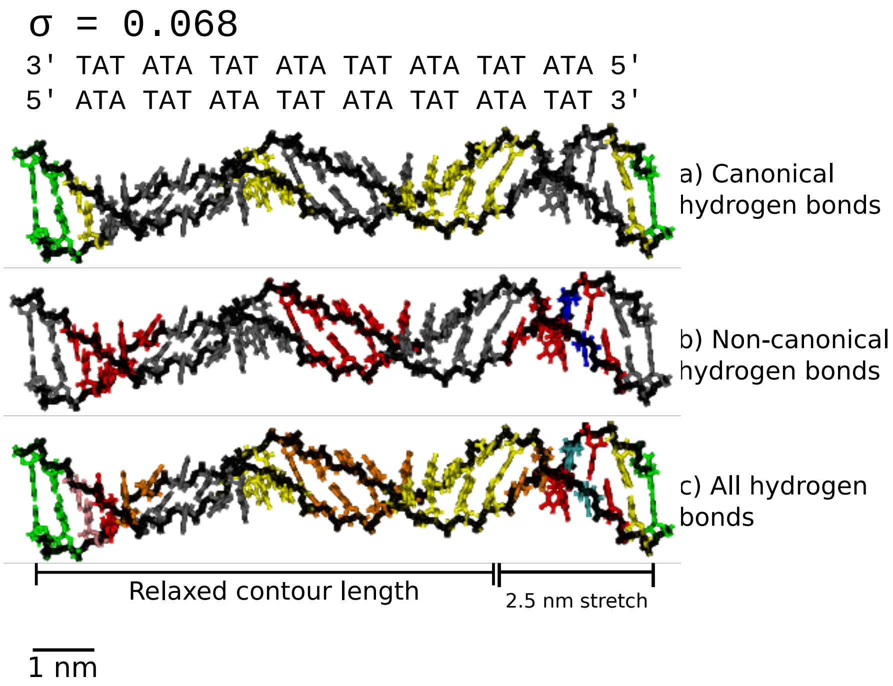


Figure 5.23: Canonical, non-canonical and all hydrogen bonds for poly-d(AT)·poly-d(AT) with $\sigma = 0.068$ stretched by 25 Å

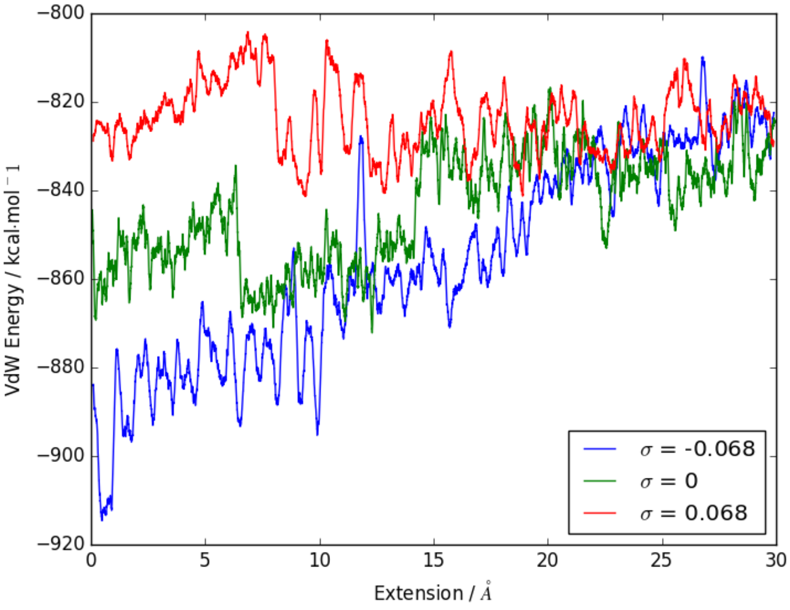


Figure 5.24: VdW energies for poly-d(AT)·poly-d(AT) with $\sigma = 0, \pm 0.068$. A 100 point rolling average has been applied.

Finally, the van der Waals (VdW) energy for the whole systems were plotted, and the results are shown in Figure 5.24. The undertwisted and untwisted DNA show the expected behaviour – the stacking interactions weakening in intensity as the duplex is stretched. However, the curve for the $\sigma = 0.068$ case is extremely interesting. The stacking interactions here stay at roughly the same level throughout the stretching, albeit with fluctuations. Once again, a graph showing the total VdW energy can not be extrapolated from. However, this is very suggestive that the DNA is engaging in more stacking interactions than hydrogen bond interactions with this perturbation applied.

In order to examine the stacking interactions of the most stable structure seen (that for which $\sigma = -0.068$), the trajectory was reanalysed and the short range stacking interaction energies were extracted. This was done using the molecular dynamics software gromacs, which has a rerun capability in which a pre-existing trajectory has energies and forces found. This analysis was done *in vacuo* as the short-range nature of the interactions renders solvation to find these energies unnecessary, and removal of the solvation scheme reduces computational cost. Rather than plot each individual frame, the frames for each stretch were averaged such that one value was found – this averaging was done after the energies had been found to ensure that no incorrect physics was found due to the inherent problems of averaging structures. The canonical and non-canonical stacking interactions were found and compared to the canonical and non-canonical hydrogen bonds, and these were compared against the average structure, as can be seen in Figure 5.25.

The stacking behaviour of the DNA is very similar to the hydrogen bonding behaviour. The canonical interactions in certain locations weaken, and in those same locations non-canonical interactions grow considerably stronger. The strengthening of the non-canonical structure is seen most in the regions where non-canonical hydrogen bonding is minimal, indicating that the structure is stabilised either by hydrogen bonding or stacking, but not both. Inspection of the average structure confirms this supposition. There are two distinct structural motifs which are repeated along the DNA contour. In the first, non-canonical hydrogen bonds dominate and the DNA base pairing retains the appearance of the canonical structure – though the pairing itself is between bases that would not usually pair in this way. In the second motif, the DNA bases are closely interlinked in the manner of a zip. Here, the stacking interactions stabilise the structure while hydrogen bonding is sacrificed. Interestingly, the stacking-stabilised structural motif does not appear with fewer than 4 bases being involved, suggesting that there is a minimum number of bases that can stably interact in this way.

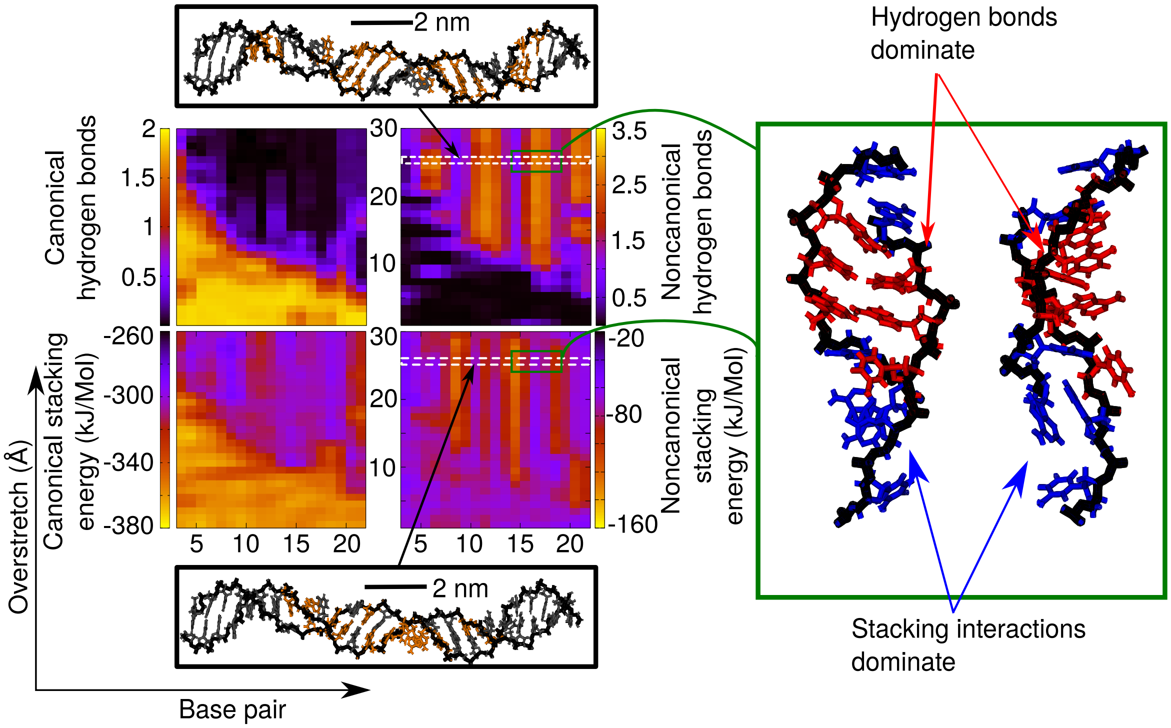


Figure 5.25: Stacking interactions and hydrogen bonding for poly-d(AT)·poly-d(AT) with $\sigma = -0.068$. The bonding energies were found using gromacs' rerun capability. The non-canonical stacking energies are seen to be strongest in regions where there is little non-canonical hydrogen bonding. Looking at the average structure and colour-coding according to the dominant interactions, it is clear that there are two coexisting structures. One is dominated by non-canonical hydrogen bonds where the bases retain an orientation similar to the canonical conformation, but with the hydrogen bonding occurring between unexpected bases. In the second structural motif the bases have interlocked in a zip-like conformation, similar to those seen in previous literature. The two motifs coexist throughout the simulation, and do not move or coalesce suggesting that this alternating conformation is highly stable and able to absorb significant structural perturbation.

That the repeating structural motifs remain in place for the duration of the simulation after forming suggests that it is a highly resilient and stable overall structure for the DNA, capable of withstanding considerable external perturbation. A repeating 4 base pair TATA unit is known to be important biologically, and is a key binding site for transcription factors. The sequence's ability to form two highly mechanically stable structural motifs may have implications for its biological role *in vivo*.

### 5.3.4 poly-d(C)·poly-d(G)

If the differences between the two AT rich sequences due to sequence are large, then the difference between AT rich and GC rich DNA is staggering. The first to be simulated was poly-d(C)·poly-d(G), and its structures at 0 and 30 Å stretch can be seen in Figure 5.26.

Even qualitatively, there are large obvious differences between this and the structures in Figures 5.12 and 5.18. The backbone in Figure 5.26 has not collapsed to anything like the degree seen previously and the bases themselves remain oriented towards the helical axis. The base pairs are somewhat inclined with respect to each other and the helical axis, which is as expected for the overstretched conformation S-DNA [93].

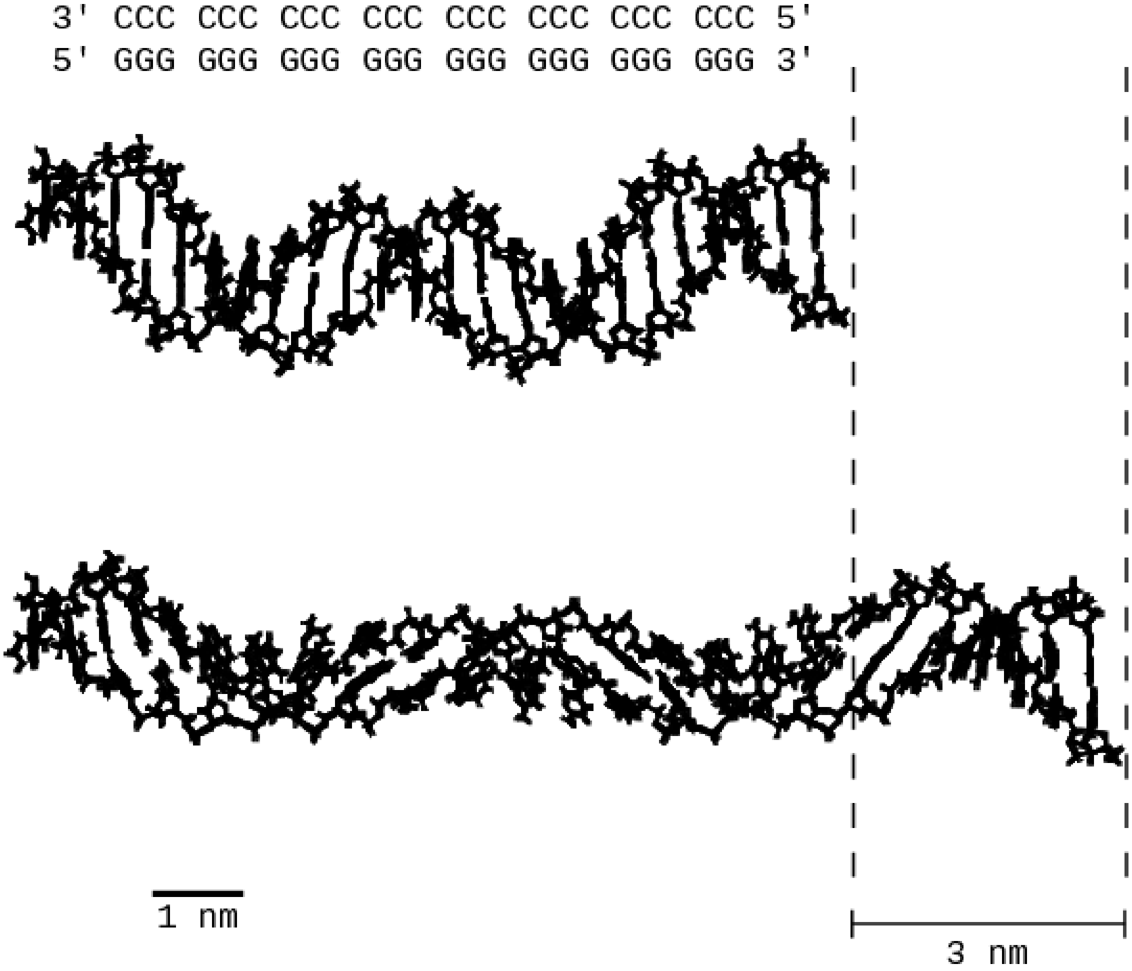


Figure 5.26: poly-d(C)·poly-d(G) with $\sigma = 0$ stretched by 0 (upper) and (lower) 30 Å

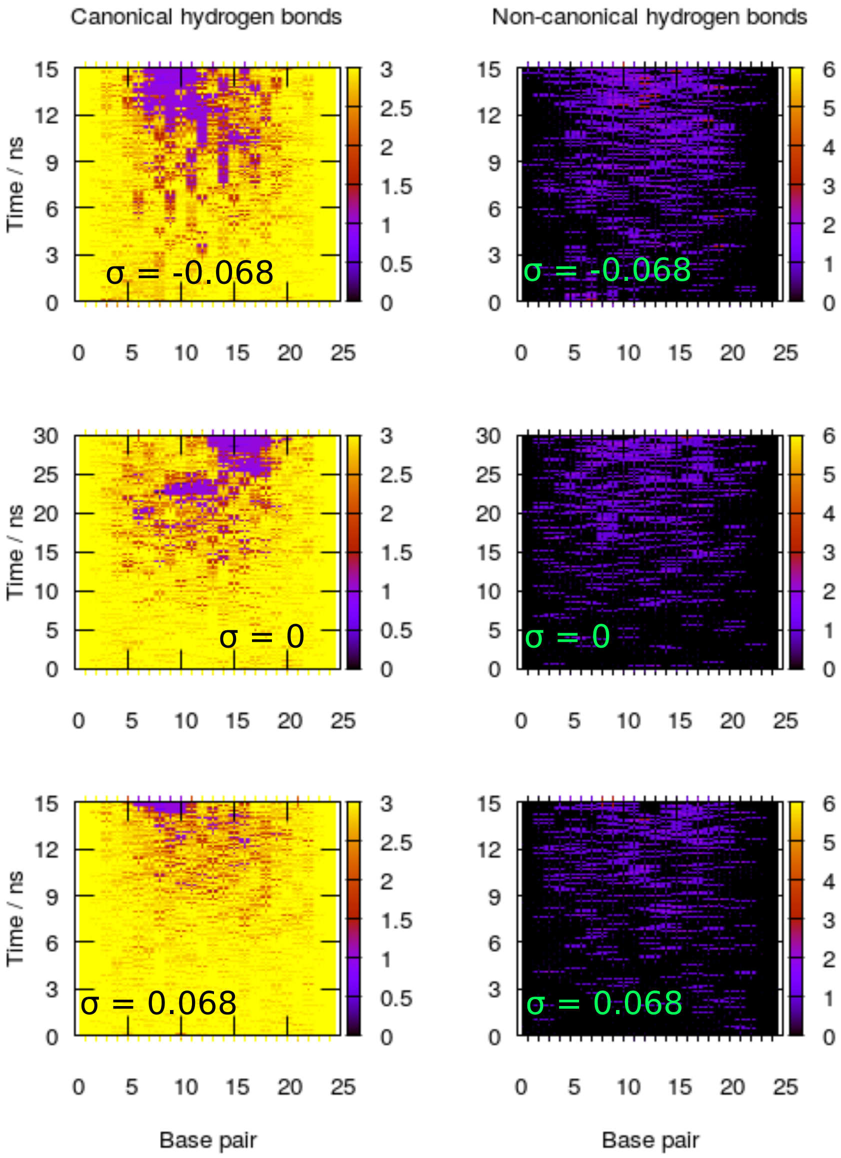


Figure 5.27: Heatmaps of canonical and non-canonical hydrogen bonds for poly-d(C)·poly-d(G) with $\sigma = 0, \pm 0.068$. On the horizontal axes is the base pair, and on the vertical axis is the simulated time. A stretch of 1 Å is applied in a single step every 500 ps for the $\sigma = \pm 0.068$ simulations and every 1 ns for the $\sigma = 0$ simulation.

To quantify how much more stable the CG rich duplex was, the heatmaps were again generated showing the canonical and non-canonical hydrogen bond interactions, and these are given in Figure 5.27. It can be seen that for the majority of the simulation, the canonical hydrogen bonds remained in place, with additional hydrogen bonds being formed presumably as the bases were pushed closer together. As the hydrogen bond analysis model is not sophisticated enough correctly to take into account three-way hydrogen bonds, it is at this stage impossible to discuss their nature definitively.

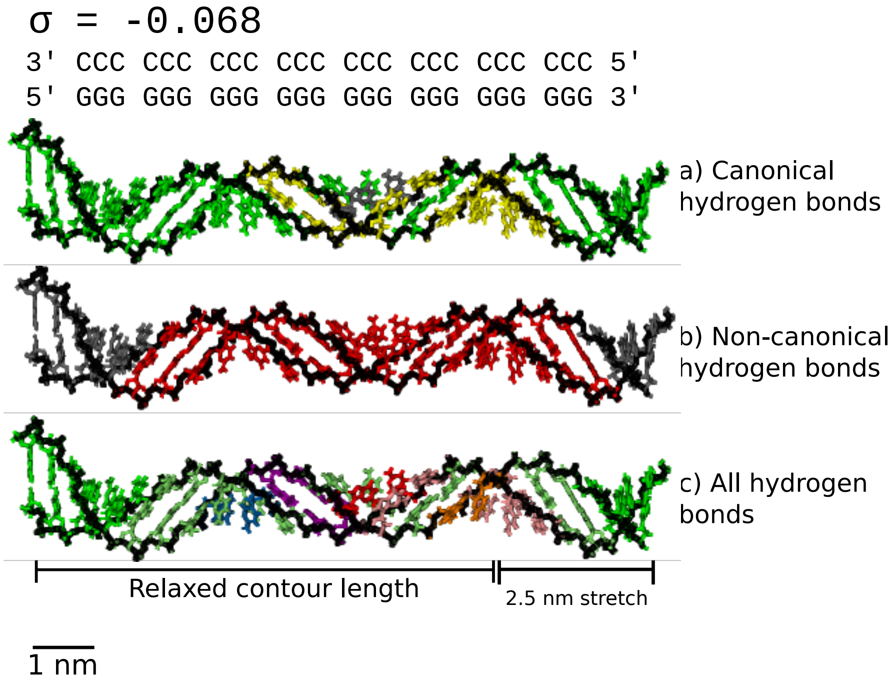


Figure 5.28: Canonical, non-canonical, and all hydrogen bonds for poly-d(C)·poly-d(G) with $\sigma = -0.068$ and extension 25 Å

By the end of the stretching, however, the duplex has begun to lose its canonical Watson-Crick pairing in the middle of the fragment (as expected, given the restrained ends), and the short lifespan non-canonical bonds grow in intensity and duration. This breakdown of the Watson-Crick base pairing occurs sooner in the undertwisted DNA, while the overtwisted DNA does not begin to lose canonical bonds until the very end of the simulation. In all cases, for much of the simulation with poly-d(C)·poly-d(G), the non-canonical hydrogen bonds are *in addition to* not *instead of* the canonical ones – something totally unseen with the previous sequences. The hydrogen bond activity has been projected on to the average structure of the duplex stretched by 25 Å and this can be seen in Figures 5.28, 5.29, and 5.30, again coloured according to Tables 5.1, 5.2, and 5.3.

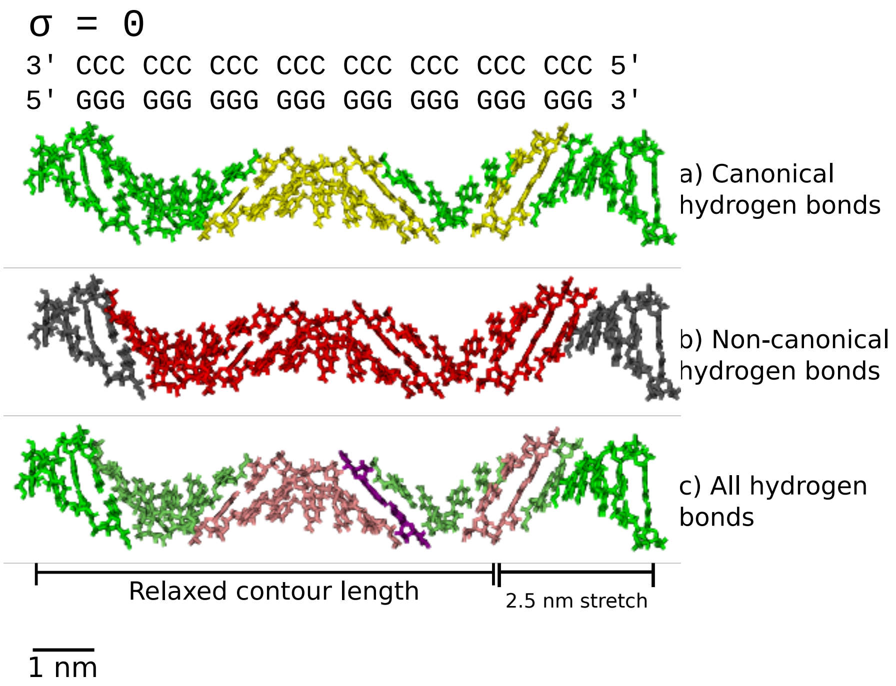


Figure 5.29: Canonical, non-canonical and all hydrogen bonds for untwisted poly-d(C)·poly-d(G) at a stretch of 25 Å

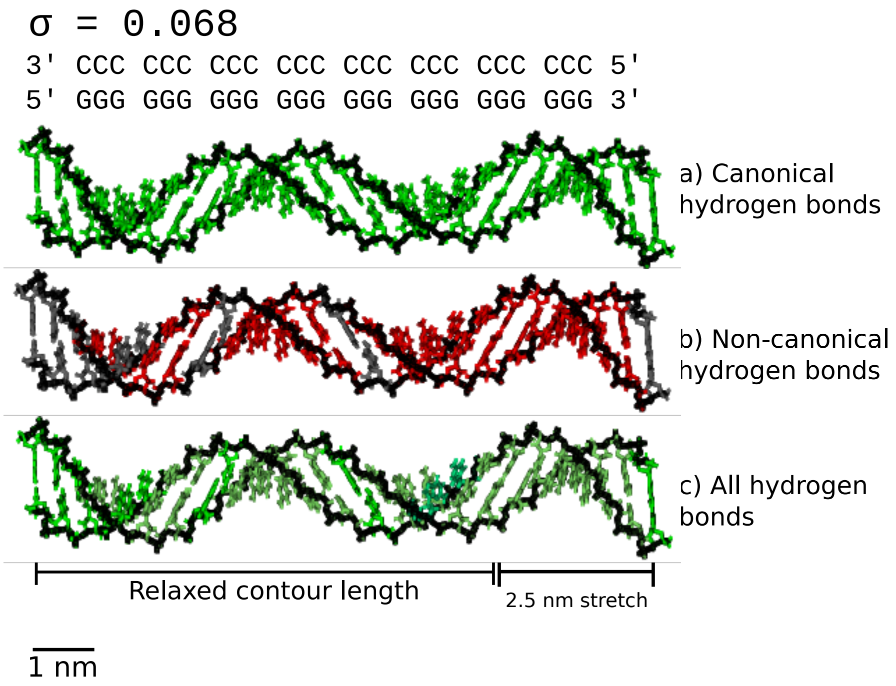


Figure 5.30: Canonical, non-canonical and all hydrogen bonds for poly-d(C)·poly-d(G) with $\sigma = 0.068$ stretched by 25 Å

These colour heat maps show again how the backbone does not collapse during stretching, and that the Watson-Crick pairing remains intact and is joined by some non-canonical hydrogen bonds. There are no grey segments of DNA (which indicates that all the base pairs are involved in hydrogen bonding), and nor are there regions where the non-canonical interactions have superseded the canonical ones. Instead, the whole structure appears compressed, and all the base pairs are interacting with not only their canonical partners but with others as well. Again, no conclusion can be drawn here about what is and is not interacting with what. However, the fact that the red is seen down the length of the DNA, instead of peaks and troughs, suggests that all base pairs are involved in some way.

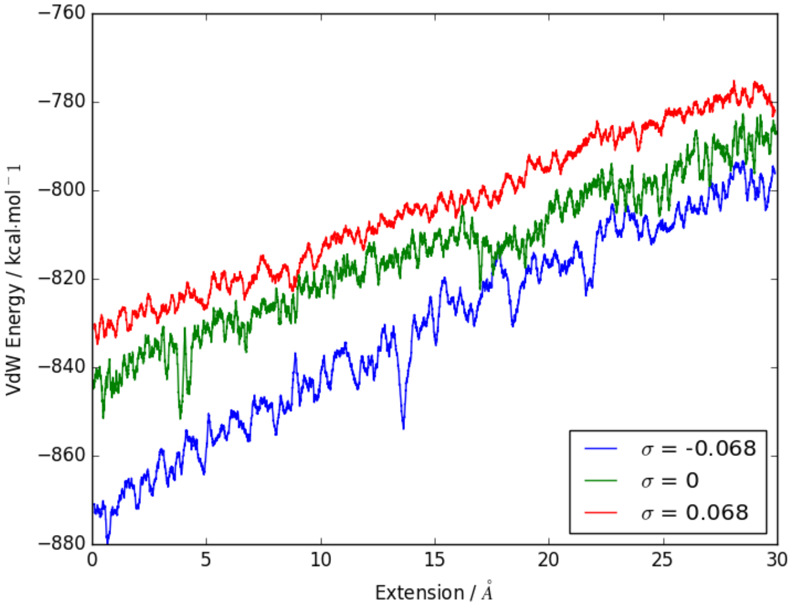


Figure 5.31: Total VdW energy for poly-d(C)·poly-d(G) with $\sigma = 0, \pm 0.068$

The total van der Waals (VdW) energy was again used as a proxy for the overall stacking interactions, and plotted in Figure 5.31. It has a very different form from those in Figures 5.24 and 5.17. Instead of high variability, in each case the VdW energy increases monotonically though noisily. This is more reminiscent of the oxDNA stretching energies seen in Figure 5.2, and is highly suggests VdW interactions stretching but not being lost or breaking and being replaced with something else. This is consistent with the structures seen above – the tilted stacking remains in Figure 5.26 for example. It appears from this graph and the structures shown above that the stacking interactions remain in place during the stretching of each twisted poly-d(C)·poly-d(G) structure. The strength of these is known to contribute to the melting temperature of DNA duplexes and so it is no surprise that they may well be involved in retaining the stability of the double helix under torsion and tension.

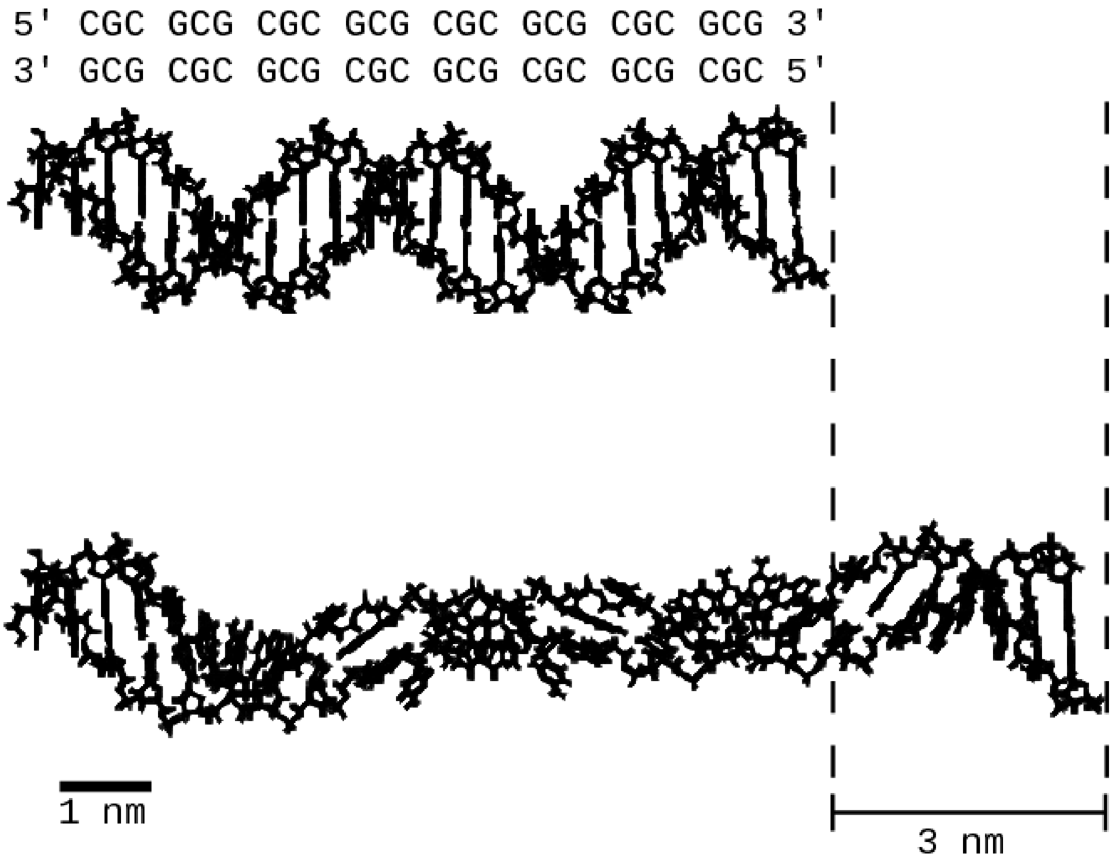


Figure 5.32: Initial (upper) and 30 Å stretched (lower) structures of poly-d(CG)·poly-d(CG) with $\sigma = 0$

### 5.3.5 poly-d(CG)·poly-d(CG)

The final sequence studied in this way with Amber was poly-d(CG)·poly-d(CG). The initial structures produced from the stretching routine (Figure 5.32) appeared some way between the extremes of the AT rich DNA and the poly-d(C)·poly-d(G). Backbone collapse can be seen, and some of the bases are sticking out away from the DNA helical axis. However, the structure is more orderly than in the A rich examples above. This initially implies that the response to twist and stretch is dependent not just on AT and GC content but on the exact sequence of the fragment in question.

The heatmaps of hydrogen bonds is shown in Figure 5.33. As with the poly-d(C)·poly-d(G), undertwisting the DNA leads to quicker loss of the Watson-Crick base pairs, while the overtwisted DNA remains the most stable over a range of extensions. The non-canonical hydrogen bonds more closely mirror the loss of canonical bonds than those in the previous case, but also exist in addition to the canonical bonds instead of immediately replacing them. The non-canonical bonds appear also to exist more transiently, with the region of non-canonical bonding moving along the DNA through time.

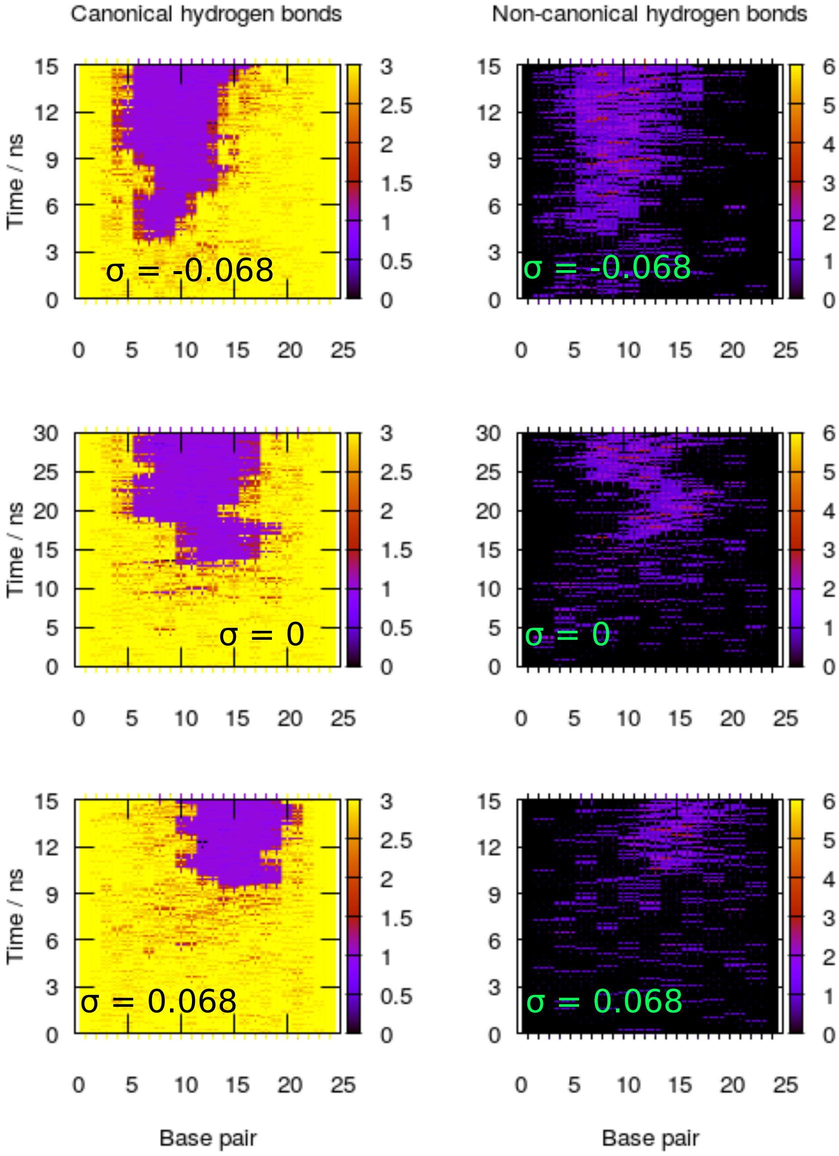


Figure 5.33: Heatmaps of canonical and non-canonical hydrogen bonds for stretched and twisted poly-d(CG)·poly-d(CG) with $\sigma = 0, \pm 0.068$. The horizontal axis shows the base pair, and along the vertical axis is the simulated time. A sharp instantaneous stretch of 1 Å is applied every 500 ps for the $\sigma = \pm 0.068$ simulations and every 1 ns for the $\sigma = 0$ simulation.

There is not much pattern in the non-canonical bonds, no strong motifs enter the DNA. The few places where more than one non-canonical bond are generated do not sustain such bonds for longer than one or at most two stretching events. Meanwhile, the solid purple of the Watson-Crick heatmaps suggest that it is preferential for the DNA to lose precisely two of its canonical hydrogen bonds in order to spatially reorient. The projection of these heatmaps on to the average structure of the 25 Å stretched poly-d(CG)·poly-d(CG) duplex are given in Figures 5.34, 5.35, and 5.36.

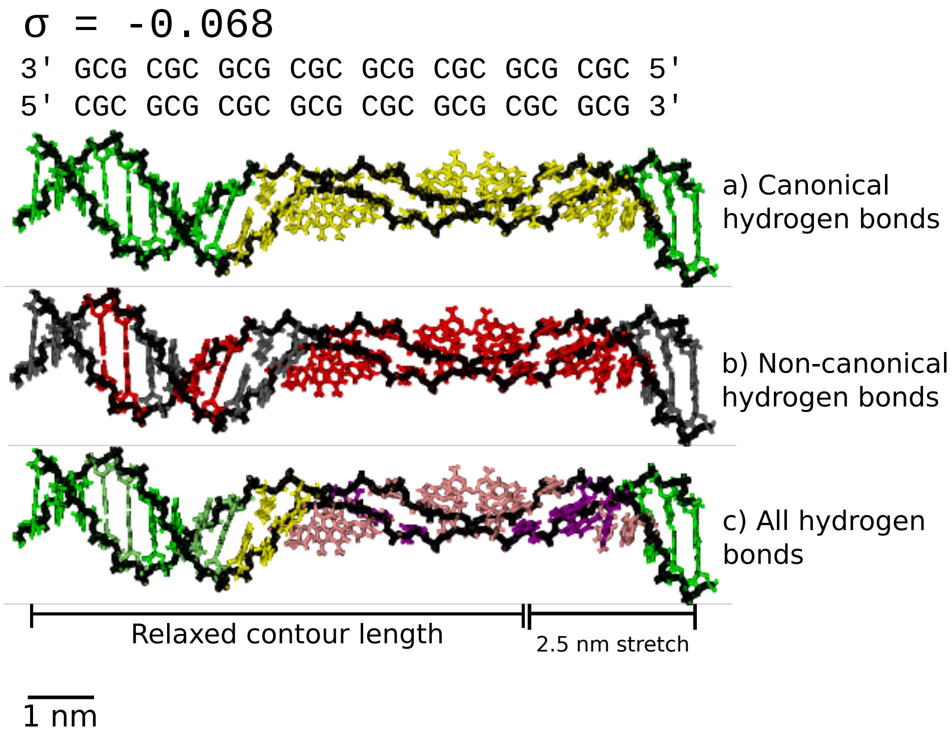


Figure 5.34: Canonical, non-canonical and all hydrogen bonds for poly-d(CG)·poly-d(CG) with $\sigma = -0.068$ and applied stretch 25 Å

It can be seen that indeed in each case the backbone has collapsed more than for poly-d(C)·poly-d(G) while retaining more hydrogen bond base pairing than in either of the AT rich sequences. The loss of Watson-Crick hydrogen bonds and increase of non-canonical bonds is largely localised to the same areas suggesting that the structure is relaxing and stabilising by collapsing its backbone and substituting some non-canonical hydrogen bonds for the conventional ones. Key however is the difference between these structures and those for poly-d(C)·poly-d(G). The differences seen between them confirms that it is not only the hydrogen bonds which are determining the structure, but the stacking interactions as well, and those interactions are sequence specific.

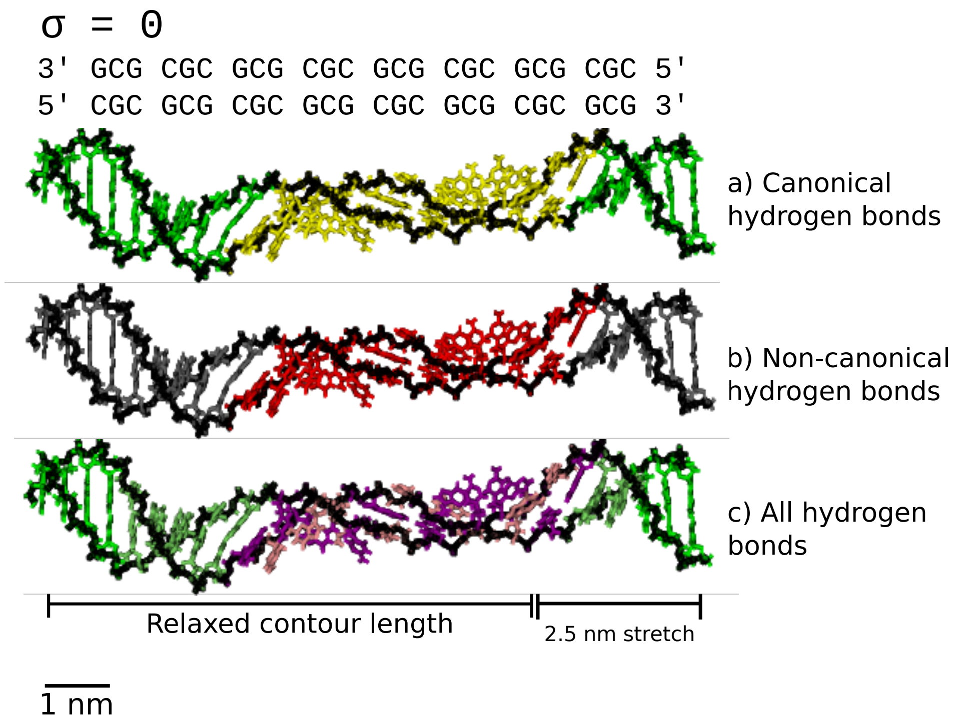


Figure 5.35: Watson-Crick, non-Watson-Crick, and all hydrogen bonds for poly-d(CG)·poly-d(CG) with $\sigma = 0$ and applied stretch 25 Å

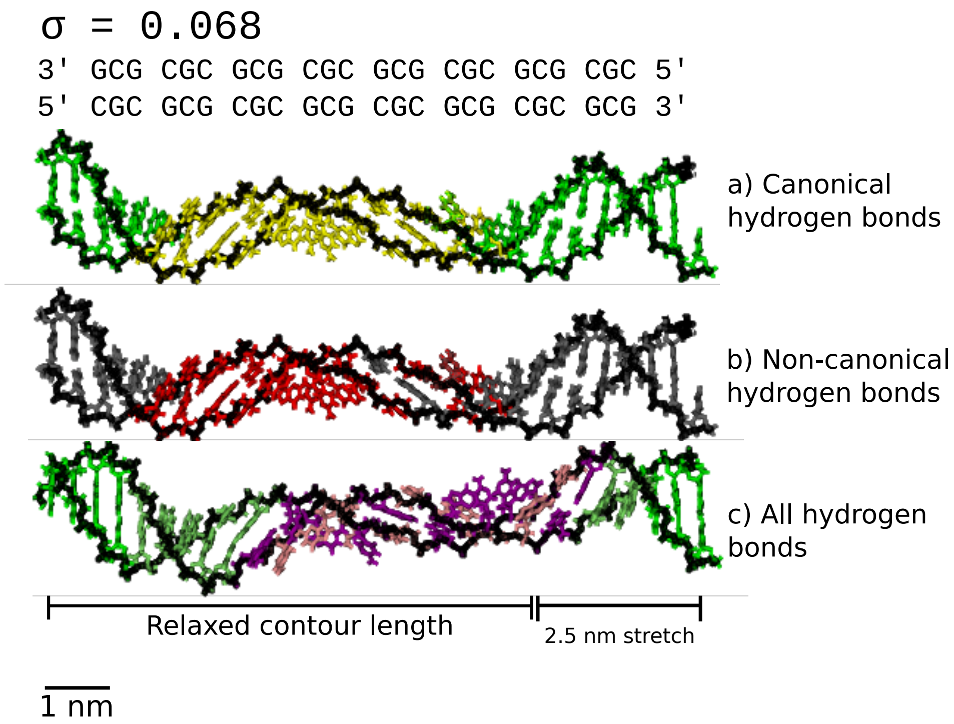


Figure 5.36: Canonical, non-canonical and all hydrogen bonds for poly-d(CG)·poly-d(CG) with $\sigma = 0.068$ and applied stretch 25 Å

This is further evidenced by the graph of VdW energy through time in Figure 5.37. An overall decrease in total VdW energy at approximately 15 Å for the untwisted duplex and at 20 Å for the overtwisted duplex correspond with the loss of hydrogen bonds as seen in Figure 5.33. Taken together with the hydrogen bond data, this suggests that stacking interactions are being favoured as a stabilising mechanism over hydrogen bonds for this particular sequence.

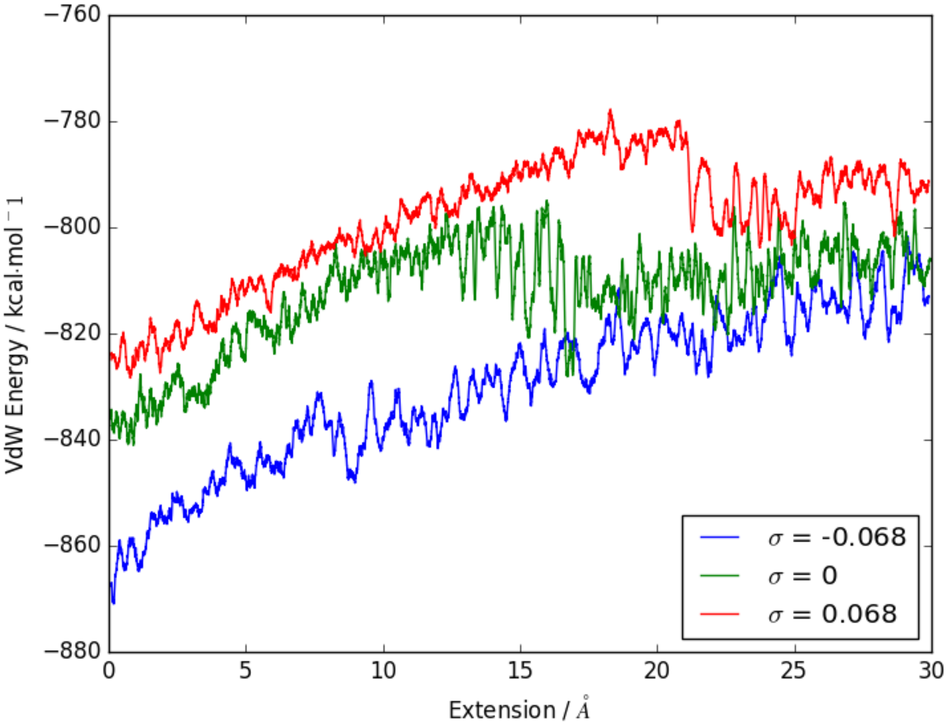


Figure 5.37: Total van der Waals energy through time for poly-d(CG)·poly-d(CG) with $\sigma = 0, \pm0.068$

## 5.4 Summary

In this Chapter, the parameters used to perform both Amber and oxDNA simulations were introduced and justified. The oxDNA simulations were shown to demonstrate localised stretching behaviours instead of a homogeneous stretch, and the backbone was seen to collapse and base pairs to become exposed consistent with previous literature. The localised stretching justified using short stretches of DNA to more closely probe the effects of sequence and applied twist to the stretching characteristics of DNA.

Four structures were chosen which together give insight into whether only the AT or GC content is important or if other factors are at work.

For poly-d(A)·poly-d(T), the duplex was seen to rapidly lose its Watson-Crick base pairings and in their place obtain non-canonical hydrogen bonds as the backbone collapsed and the bases turned outwards. poly-d(AT)·poly-d(AT) was shown to have similar qualitative behaviour, but the undertwisted structure formed a regular repeating unit of increasing and decreasing numbers of non-canonical hydrogen bonds, as – to a lesser extent – did the untwisted sequence, indicating a new structure was being formed and was stable over a range of extensions. In each case, the non-canonical bonds formed where canonical ones had broken, replacing the Watson-Crick interactions entirely after a certain amount of applied stretch. In both the AT-based sequences, the stacking interactions as reported by the van der Waals (VdW) energies was noisy and irregular, and did not correspond obviously to any changes in hydrogen bond conformation.

poly-d(C)·poly-d(G) was the most stable of all the sequences interrogated here. It maintained its canonical base pairing for almost the full simulation in the case of the overtwisted structure. The base pairs were seen to incline with respect to the helical axis as is expected from the stretched form of DNA known as the S-form [50], and the non-canonical hydrogen bonds which were seen were in general alongside the canonical ones instead of replacing them. More structural instability was seen when undertwisting this sequence, indicating that the tight packing of bases in the untwisted and overtwisted structures may form a part of their stability. The VdW energy rose monotonically with stretch indicating that the stacking interactions were weakened but not destroyed, and there was no evidence of new stacking interactions being formed.

This was not the case for poly-d(CG)·poly-d(CG). The backbone collapsed similarly to that of the AT rich DNA, and the bases lost their canonical hydrogen bonds in favour of non-canonical ones. The non-canonical bonds forming corresponded again to the loss of the traditional Watson-Crick hydrogen bonds, and this was also correlated in the overtwisted and untwisted structures with a drop in VdW energy. This is highly suggestive of hydrogen bonds being sacrificed in favour of stronger stacking interactions, suggesting a formation of a structure of interlinked bases interacting through the VdW stacking interactions as in [1].

All of this taken together shows DNA stretching is not homogeneous and the that the specific response is not only a function of AT or GC content. Instead, exact sequence and the precise nature of the applied perturbation can lead to new stable structures forming, and the sequence itself may determine whether stacking effects or new non-canonical hydrogen bond formation is preferable. In order to more rigorously understand this, further quantitative investigation is needed to assess the distribution of both canonical and non-canonical stacking interactions in the sequences studied.

This data sheds initial light on the structural changes and states of DNA in the early stages of overstretching, before and around the overstretching transition. The response of different sequences to even moderate stretches is striking enough that there are implications for molecular machinery, and the resilience of different sequences to stretching and twisting may have impact on DNA-protein interactions as well as packaging, replication, and repair.

With the data presented in this Chapter, predictions may also be made about the fluorescence response of overstretched and torsionally constrained DNA. Structures which have kept their canonical hydrogen bonds are likely to be more favourable for the binding of an intercalating dye like YOYO-1, and those whose backbones collapse are more likely to exhibit loss of fluorescence intensity. This is true also for groove binding dyes – if the groove no longer exists for the dye to bind to, the DNA will not be visible. With a known sequence in the 15 kbp DNA tether, an obvious next step would be to apply similar forces and twists and use fluorescence microscopy to examine the distribution of dyes to compare to this prediction.

# Chapter 6

# Discussion

The aim of this project was to develop methods for use on a bespoke magneto-optical tweezing microscope in order to apply forces and torques to individual strands of DNA, and to better understand the mechanics of DNA stretching under torsional constraints through multiscale molecular dynamics simulations.

Experimentally, the project can be considered successful in that it is now possible to form DNA tethers and manipulate them with magneto-optical tweezers while extracting force data that is in agreement with other work and close to theoretical expectations. Differences between labelled and unlabelled DNA were seen in Chapter 3, especially in the different buckling severities. The persistence length was found to be in the correct range, though a reliable population average of persistence lengths could not be obtained. However, the contour length was consistently higher than the expected value calculated from the canonical B form.

This could be for two reasons. Analysis of the data collected could have gone wrong, or the system itself does have a conformation relatively far from the canonical structure. The analysis itself is done in with well-established tools (Python and SciPy), and using well-established models like the wormlike chain [33]. Although such fitting is sensitive to initial conditions, it is unlikely that error-prone fits would consistently get the same wrong answer, and especially the same wrong answer where the values of persistence lengths found are in the correct range. Further, constraining fits to have $4\,\mu\text{m} < L_C < 5\mu\text{m}$ resulted in extremely poor fits. Instead, it seems probable that buffer conditions are the reason for the values found. The buffer used is a complex recipe, and its effects on dsDNA could include an increase of the overall contour length. The temperature in the vicinity of the magnetic bead is also not precisely known, and as this is a term in the wormlike chain model it could have an effect. However from experimentation with various values the author believes the temperature not to be the key factor.

The increase in contour length upon binding to YOYO-1 is known to be around 38%, yet in this work no such increase was seen. Once again, the buffer conditions or damage to the DNA itself could result in a more open conformation of DNA and hence a longer contour length than would be expected from the standard approximation of 0.3 nm/bp. The fluorescence assays were shown to work to an extent – the DNA is visible. However, super-resolution information is not yet available as the damage to the DNA was too great, and the background brightness too high. Further optimisation of this is clearly needed.

The instrument, as it currently stands, is nonetheless a useful tool for interrogating DNA dynamics in brightfield. The QPD sensitivity and nanostage precision were found to be very good, and the force clamp works as it should as shown in Chapter 2. The expected behaviour was found qualitatively for rotated DNA held under both high and low tension, and the value at which the undertwisted DNA did not shorten was found to be between 0.55 and 0.7 pN, in excellent agreement with previous work which found a range of 0.6-0.8 pN[92]. However the buffer conditions in the two works were not the same, and a repeat of the experiment in 10 mM phosphate buffer as in [92, 144] for a rigorous comparison.

It was not possible in this project to look for the overstretching transition at 60 pN – due to the optical properties of the magnetic bead, that much force cannot be exerted. With smaller, nonmagnetic beads, such forces would be possible. This is a strength – in that it is physiological forces which are examined – but it is a weakness in that it is a limit and nothing above around 2 pN can be compared. As physiological forces are often on the order of single pN (for example, the molecular machines myosin and kinesin have stall forces <10 pN [166]) it is generally preferable to compromise on the absolute force applied in order to look at force and torque at the same time. That being said, it would be a good idea in the future to find a different bead or re-design the trap such that the forces can go higher by at least an order of magnitude in order to examine the DNA overstretching regime and to be able to apply higher forces to more powerful molecular machines such as RNA polymerase which has a stall force around 20 pN [166]. Alongside this, adding in a spatial light modulator or acousto-optical deflector would allow movement of the trap and force clamping by trap movement not nanostage movement, and dual trap creation, which may be a useful feature in future as they would allow nucleic acid dumbbell experiments to be performed.

Most interestingly from the experimental work, significant hysteresis behaviour was seen when rotating and relaxing DNA tethers under a constant force. It is contentious in the literature whether such hysteresis exists, and whether it is observed is likely to be a function of how close to thermodynamic equilibrium the experiments take place at, as well as considerations such as force clamp response speed, and applied tension. With rotations at 2 Hz and tensions around 1 pN, the experiments in this thesis showed clear hysteresis-like behaviour and thus were likely far from equilibrium. This presents an opportunity to understand mechanical properties of DNA which have not so far been interrogated.

Although the magneto-optical tweezers have been shown to work in the sense of creating, stretching, and twisting tethers, and gathering data in agreement with previously published results, technical challenges remain. For example, the tethers should be formed perfectly along the nanostage oscillation axis. To do this the LabVIEW code should be updated such that it puts the tether under tension in one axis and minimises the force in the other.

A primary challenge however is the low throughput of the tweezers. In common with many other groups (anecdotally) a day is successful if there is more than one usable tether formed. Because of this low throughput, little fluorescence optimisation could be done. The gathering of data in brightfield took the majority of the time after the method development.

For this reason, the most important addition to the microscope at this stage would be a microfluidics system. This would allow protein to be added while monitoring a tether, would allow for more flexibility with slide treatment and buffer changes, and would make possible long, multistage assays observing one tether under different conditions. With microfluidics in place, the low throughput may also be mitigated: conditions in the DNA binding buffer could be optimised such that binding happens more repeatably, and then the imaging buffer could be substituted in with microfluidics.

The surface adhesion with nitrocellulose is relatively reliable but the large beads can come loose from the surface, and magnetic beads and DNA can adsorb to the surface despite the BSA blocking. A specific interaction between the antiDIG beads and a BSA/digoxigenin treated surface could be preferable but would take significant time to optimise and so could not be done as part of this project.

In the theoretical work, significant novel findings were obtained from the molecular dynamics. In Chapter 5 it was shown that a small overstretch applied to large piece of DNA results in a largely local perturbation. Therefore, a small fragment of 24 bp was sufficient to capture the initial stages of overstretching. The DNA response was found to be dependent on applied force, torque, and specific base pair sequence, rather than just AT/CG content. The poly-d(A)·poly-d(T) and poly-d(AT)·poly-d(AT) exhibited very different behaviour, with the AT repeating unit forming a stable, repeating structural unit stabilised by non-canonical hydrogen bonds in certain regions and non-canonical stacking in others. The two modes of stabilisation were shown to be separate, and the repeating motifs appeared to be highly mechanically stable, remaining in place throughout the simulation after they had initially formed.

By contrast, the AA repeating unit, simply melted. This is evidence that the current approximations used – that it is simply AT/CG content which determines DNA structural behaviour under thermodynamic and mechanical processes – are too coarse at this length scale. In order to predict experimental behaviour well, it is necessary to take into account the localisation of the perturbation and the exact sequence that appears there.

The CC and CG repeating units were found to have largely similar responses under torsionally constrained stretching, with each forming a conformation similar to S-DNA that is characteristic of the freely rotating stretched DNA seen in other simulations [50]. These simulations suggest that it is possible to predict the disrupted regions in a slightly overstretched torsionally fixed dsDNA duplex, based on sequence and position in the construct. The response found experimentally to such disruption would likely be fluorescence drop-off in those regions, and with a refined superresolution methodology these dark regions could be pinpointed to within a few tens of base pairs – justifying or contradicting these simulations.

## 6.1 Future Work

The life of this microscope is really just beginning. The number of different assays possible with rotation and force application are enormous as shown by the volume of literature in Chapter 1. However, the microscope must walk before it runs. The most important experiment to carry out now will be the control experiment in buffers previously reported. The buffer used was a sensible option when trying to optimise the fluorescence, but it is time now to perform the control experiments that will pinpoint the accuracy of the machine. It has been shown that the force value above which undertwisted DNA will not shorten is in the range of 0.55-0.7 pN, but this should be found more precisely and compared with known values. This should be done with the force in the perpendicular axis minimised to obviate the need to rescale forces and displacements and these experiments would be the first ones the author would perform with longer time in the project.

Once that is done, the fluorescence should be optimised next. Preliminary dye concentrations and laser power settings have been found as part of this work and should be considered a starting point to work from. However, they will need further refinement. The author suspects that the DNA should be complexed with the YOYO-1 and purified as in Chapter 2, but then a small concentration of YOYO-1 should be present in the imaging buffer as well to replace any that unbinds stochastically.

The scavenging buffer used appears to work but a rigorous titration of each component in turn would probably increase its efficacy. Even without such a titration, more time optimising the YOYO-1 conditions would certainly improve the fluorescence images and improve the utility of the microscope overall.

With the forces further validated by controls, brightfield experiments looking specifically at the hysteresis behaviour should be performed also. Experiments performed in the same way as in this thesis but in a more pared down buffer such as the previously used 10 mM phosphate buffer [92, 144], with rotation rates varying from 0.1 to 2 Hz, would verify if the effect is dependent on buffer, rotation speed, or both. The author speculates that less apparent hysteresis will be seen with lower ionic strength buffers (as less electrostatic screening will occur so the DNA will not be able to access so many conformations) and with decreasing rotation speed, as the whole process will be closer to thermodynamic equilibrium. The constant rotation speed assays could also be used to interrogate the effect of rotation on molecular machines in a similar way to the experiments done on force inhibition of polymerase.

Summing this up, a roadmap for a notional PhD student or postdoc carrying on this work could be:

1. Adapt LabVIEW code to minimise force along the non-tether axis
2. Repeat brightfield experiments in Chapter 3 in low ionic strength conditions
3. Repeat the hysteresis experiments with rotation rates such as 0.1 Hz, 0.5 Hz, 1 Hz, 5 Hz
4. Optimise the YOYO-1 by running a titration of YOYO-1 added to the final imaging buffer
5. Rerun the rotation experiments in fluorescence and observe plectoneme formation
6. Overstretch a tether in fluorescence and attempt to correlate with the simulation data in Chapter 5
7. Do step 6 at multiple $\sigma$

8. Over and undertwist DNA in fluorescence
9. Add microfluidics equipment to the microscope

This would be an interesting project in its own right (especially stages 3 and 6) and would show significant new images of DNA conformational change. The microscope would also then be ideally suited for use with topoisomerases or other molecular machines, and could correlate the fluorescence loss in the vicinity of the molecular machine with force and torque transduced, both by measurement of the bead movement and by comparing with known control experiments and simulations.

Finally, the simulation work could be extended to fully characterise the molecular dynamics trajectories found, for instance by going beyond the VdW energies and using quantum mechanics to examine the $\pi/\pi*$ orbital interactions which are ultimately responsible for the DNA base stacking.

Ultimately, the future work should ideally link the two regimes examined in this thesis. It is unfortunate that the forces and torques applied to the DNA *in vitro* and *in silico* are not comparable at present, due to limits on computer time and current limits on the forces applied to the magnetic bead, but these hurdles can be overcome with different beads, a higher laser power, and further simulations.

If the roadmap above were followed the future use should be to concentrate on biomedical issues. Specifically, drug binding to DNA could be well characterised, as could the inhibition of molecular machine activity by drug addition. The reversibility or otherwise of this process could be examined using microfluidics to wash the drugs in and out, for example. The ultimate goal would be a microscope which could fluorescently image DNA, a molecular machine, and the drug effect on the system all while measuring forces and torques experienced by the DNA. This would open up a new avenue for drug creation and testing, as well as shedding light on the fundamental processes of life.

# Appendix A

# Protocol for Constructing the 4 kbp Tether

Here is reproduced the full protocol for synthesising the 4 kbp DNA construct, taken from Dr Zhou's PhD thesis.

## Step 1

Procure the following primers from a commercial supplier:

| Name | mass (mg) | amount (nmol) | sequence |
|---|---|---|---|
| pET28a forward | 8.14 | 739.3 | 5'- GGA CCT GCT TTC CAA CGC CAT ATT CAA CGG GAA ACG - 3' |
| pET28a reverse | 6.90 | 707.6 | 5'- GGG TCT CGA CCA AAC AGC TGA TTG CCC TTC AC -3' |
| pBluescript forward | 7.65 | 688.5 | 5'- GGA CCT GCT TTC GTT GTG GCG TAA TCA TGG TCA TAG -3' |
| pBluescript reverse | 6.96 | 704.7 | 5'- GGG TCT CGT GGT TTA TAG TCC TGT CGG GTT TC -3' |

and add 1000 µL double distilled water to each tube to obtain:

| Name | amount (nmol) | added water ($\mu$L) | Final concentration ($\mu$M) |
|---|---|---|---|
| pET28a forward | 739.3 | 1000 | 739 |
| pET28a reverse | 707.6 | 1000 | 708 |
| pBluescript forward | 688.5 | 1000 | 689 |
| pBluescript reverse | 704.7 | 1000 | 705 |

## Step 2

Further dilute the primers by taking 1 µL primer solution from each tube and adding 9 µL water to each solution:

| Name | vol ($\mu$L) | added water ($\mu$L) | Final concentration ($\mu$M) |
|---|---|---|---|
| pET28a forward | 1 | 9 | 73.9 |
| pET28a reverse | 1 | 9 | 70.8 |
| pBluescript forward | 1 | 9 | 68.9 |
| pBluescript reverse | 1 | 9 | 70.5 |

## Step 3

Mix the following chemicals in a PCR tube for PCR reaction to generate the central part. Note that GoTaq Green (2x) is composed of 400 $\mu$M dATP, 400 $\mu$M dGTP, 400 $\mu$M dCTP and 400 $\mu$M dTTP so the amount of dNTP is 400 $\mu$M × 12.5 $\mu$L × 4 = 20000 pmol. Central part is 6254 nt so 20000 pmol dNTP has the potential to make 20000/6254 = 3.2 pmol central parts. pET28a is measured with a NanoDrop 2000/2000c Spectrophotometer to have initial concentration of 92.6 ng/ul = 14.2 nM. It is known that the spectrophotometer may have significant error so the measurement is only a reference. Initial concentration means concentration of reagents before mixing; final concentration means that after mixing.

| Chemical | Volume ($\mu$L) | Initial conc ($\mu$M) | Final conc ($\mu$M) | amount (pmol) |
|---|---|---|---|---|
| GoTaq Green 2x | 12.5 | 2× | 1× | - |
| pET28a forward | 0.5 | 73.9 | 1.48 | 37 |
| pET28a reverse | 0.5 | 70.8 | 1.42 | 35.5 |
| pET28a | 3 | $1.42 \times 10^{-3}$ | $1.7 \times 10^{-3}$ | $42.6 \times 10^{-3}$ |
| Water | 8.5 | - | - | - |

## Step 4

Mix the following in a PCR tube for PCR reaction to generate the digoxigenin functionalised handles. The grey texts are components of the reagent in the line above them. It is necessary to write out each dNTP both for GoTaq Green and for the labelling reagent to calculate the number of dig on each handle. The pBluescript above is measured with a NanoDrop 2000/2000c Spectrophotometer to have initial concentration of 486 ng/$\mu$l = 135 nM. So the final concentration is 0.5/25 × initial conc = 9.72 ng/$\mu$L

= 2.7 nM. The total dNTP is 30000, the handles are 1050 nt long, so the reaction can make 30000/1050 = 28.6 pmol handles.

| Chemical | Vol ($\mu$L) | Initial conc ($\mu$M) | Final conc ($\mu$M) | Amount (pmol) |
|---|---|---|---|---|
| GoTag Green 2× | 12.5 | 2× | 1× | - |
| DIG labelling mix 10× | 2.5 | 10× | 1× | - |
| pBluescript forward | 0.5 | 68.9 | 1.38 | 34 |
| pBluescript reverse | 0.5 | 70.5 | 1.41 | 35 |
| pBluescript | 0.5 | $135 \times 10^{-3}$ | $2.7 \times 10^{-3}$ | $68 \times 10^{-3}$ |
| Water | 8.5 | - | - | - |

## Step 5

Mix the following in a PCR tube for PCR reaction to generate the biotin functionalised handle. The effective amount of dNTP for each species is 5000 pmol, so can make 5000 × 4/1050 = 19.0 pmol products.

| chemical | vol ($\mu$L) | initial conc ($\mu$M) | final conc ($\mu$M) | amount (pmol) |
|---|---|---|---|---|
| GoTaq Green 2× | 12.5 | 2× | 1× | - |
| Biotin label 10× | 2.5 | 10× | 1× | - |
| pBluescript forward | 0.5 | 68.9 | 1.38 | 34 |
| pBluescript reverse | 0.5 | 70.5 | 1.41 | 35 |
| pBluescript | 0.5 | $135 \times 10^{-3}$ | $2.7 \times 10^{-3}$ | $68 \times 10^{-3}$ |
| Water | 8.5 | - | - | - |

## Step 6

Perform PCR reaction to create and amplify the wanted linear parts of the plasmids. Blue font means the 3 steps as a whole are repeated 30 times.

| Step | Temperature (℃) | Time | repeats |
|---|---|---|---|
| Initial melting | 98 | 2 min | 1 |
| Melting | 98 | 15 s | 30 |
| Annealing | 60 | 30 s | |
| Elongation | 72 | 5 min | |
| Final elongation | 72 | 15 min | 1 |

## Step 7

Clean up the three PCR reactions using QIAquick PCR purification kit (QIAGEN, Cat No./ID: 28104). This step removes the primers but not pET28a or pBluescript as they are too big for the purification kit to remove. The reagents created so far are summarised below. The amount of all parts have been discounted by 5% to take into account the loss during purification. The volume is set by the last step of the purification.

| DNA | amount (pmol) | vol ($\mu$L) | conc (nM) |
|---|---|---|---|
| central part | 3.04 | 50 | 61 |
| dig-handle | 27.2 | 50 | 543 |
| biotin-handle | 18.1 | 50 | 361 |

## Step 8

Restriction digestion to cut out the overhang for each part. Mix the following in three Eppendorf tubes. Incubate at 50℃ overnight. The restriction enzymes are BfuAI (R0701S), BsaI (R0535S) and the buffer is NEBuffer 3.1 (R7203S), all of which are procured from NEB.

| | vol ($\mu$L) | | vol ($\mu$L) | | vol ($\mu$L) |
|---|---|---|---|---|---|
| central part | 40 | dig handle | 40 | biotin handle | 40 |
| BfuAI | 1 | BfuAI | 0 | BfuAI | 1 |
| BsaI | 1 | BsaI | 1 | BsaI | 0 |
| 10× NEBuffer | 5 | 10× NEBuffer | 5 | 10× NEBuffer | 5 |
| Water | 8 | Water | 9 | Water | 9 |

## Step 9

Purification of the three reagents following the same procedures as in step 7. Below is calculated and measured concentration. 5% discount is applied to the calculation again to account for purification loss. The table also compares the two concentrations and finds that calculated concentrations are about 10 times higher than measured concentrations. I theorize that this is due to the inaccurate NanoDrop 2000/2000c Spectrophotometer concentration measurement (measured twice) as well as loss of reagents down the pipeline.

| | Measured conc (ng$\cdot\mu$L$^{-1}$) | Calculated conc(ng$\cdot\mu$L$^{-1}$) | Ratio |
|---|---|---|---|
| central part | 18 | 174 | 9.7 |
| dig handle | 17.7 | 273 | 15 |
| biotin handle | 13.8 | 192 | 14 |

## Step 10

Ligation. Mix the following. Incubate at room temperature overnight. This is the final product. The product is not purified because the presence of ligase helps repairing nicks formed in the DNA strands, which will be useful for the DNA to stay torsionally constrainable. For the amount and concentration of parts, measured values from step 9 are used rather than calculated values. As the construct concentration will be limited by the central part, which is 0.72 nM.

| | vol ($\mu$L) | amount ($\times 10^{-3}$ pmol) | conc (nM) |
|---|---|---|---|
| central part | 3 | 14.3 | 0.72 |
| dig handle | 3 | 80.5 | 4.02 |
| biotin handle | 3 | 59.1 | 2.96 |
| T4 DNA ligase | 1 | - | - |
| 10× DNA ligase buffer | 2 | - | 1× |
| Water | 8 | - | - |

# Appendix B

# Heatmaps for 100 ps/Å Replicas

Here, one to a page, are presented the canonical and non-canonical hydrogen bonds per base pair with stretch.

## B.1 poly-d(A)·poly-d(T)

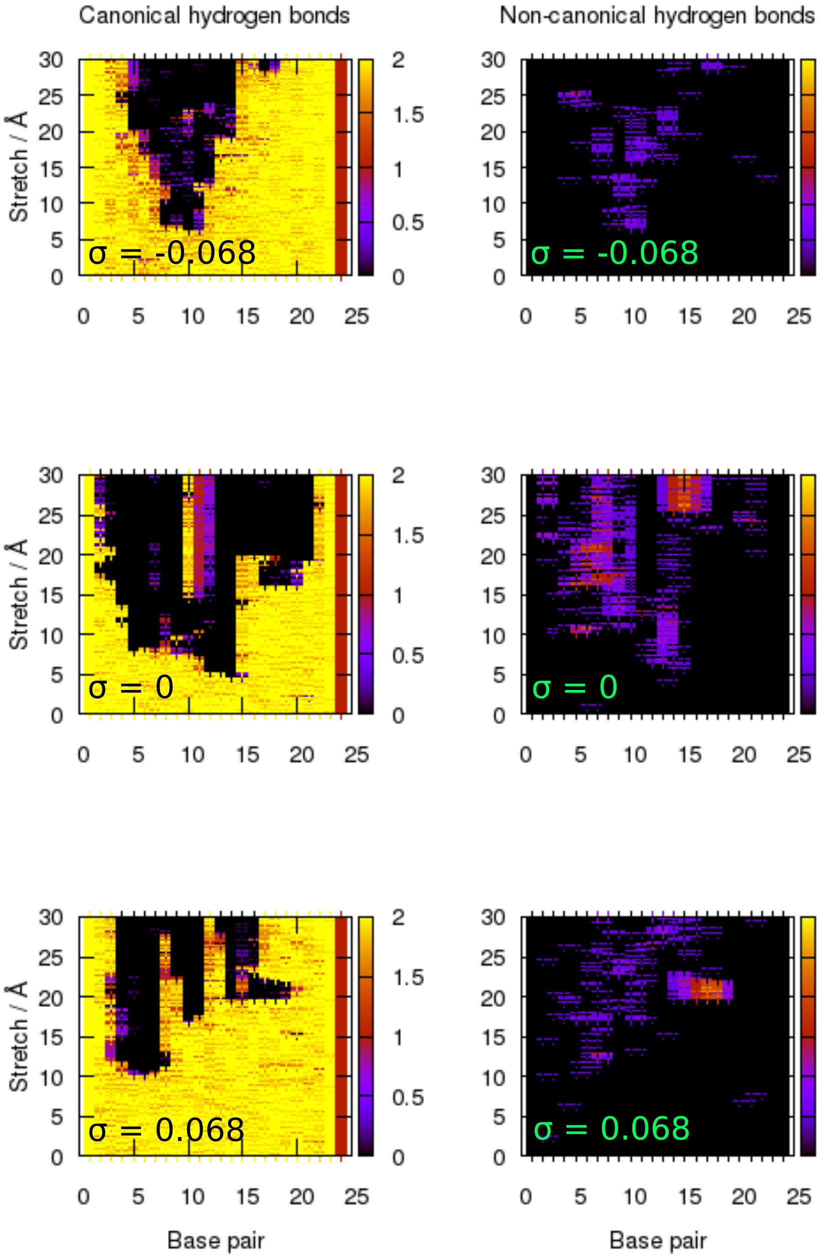


Figure B.1: poly-d(A)·poly-d(T) 100 ps/Å replica 1

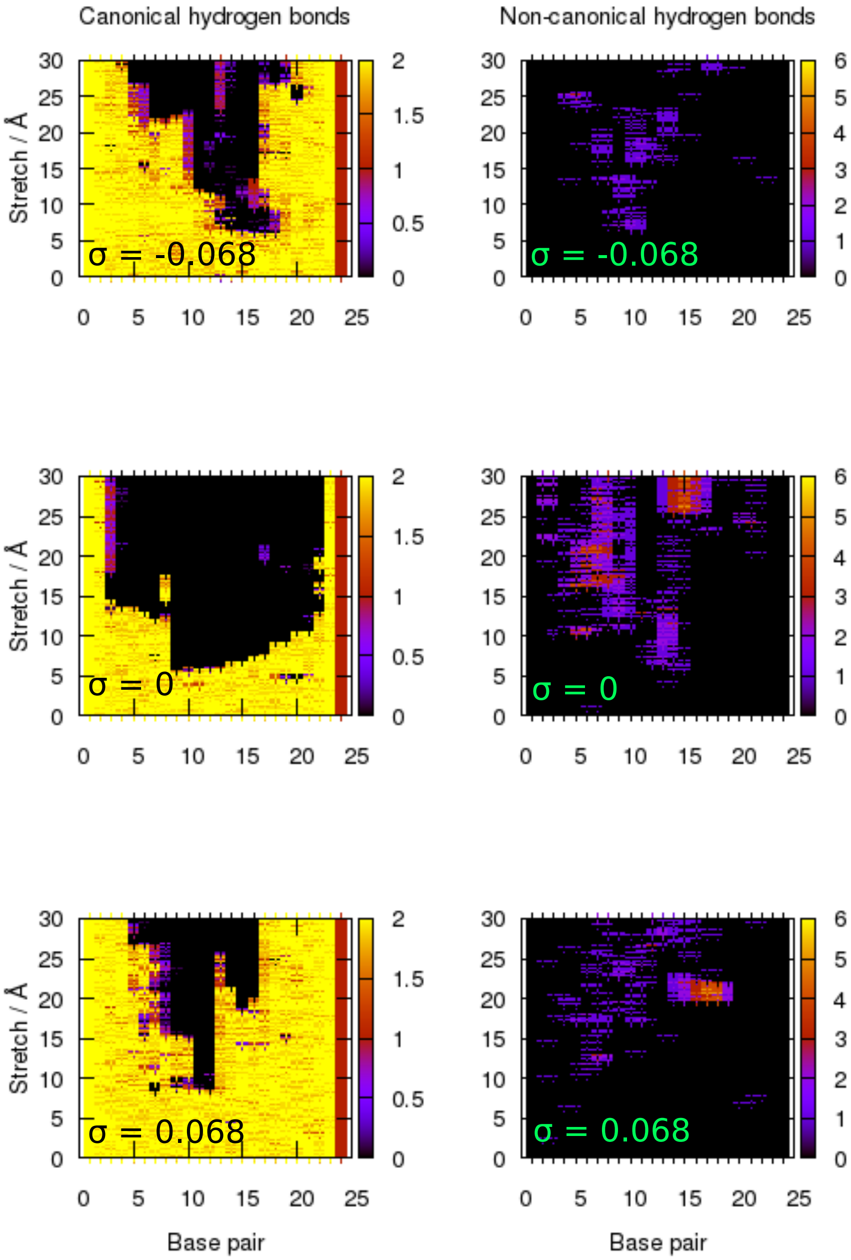


Figure B.2: poly-d(A)·poly-d(T) 100 ps/Å replica 2

## B.2 poly-d(AT)·poly-d(AT)

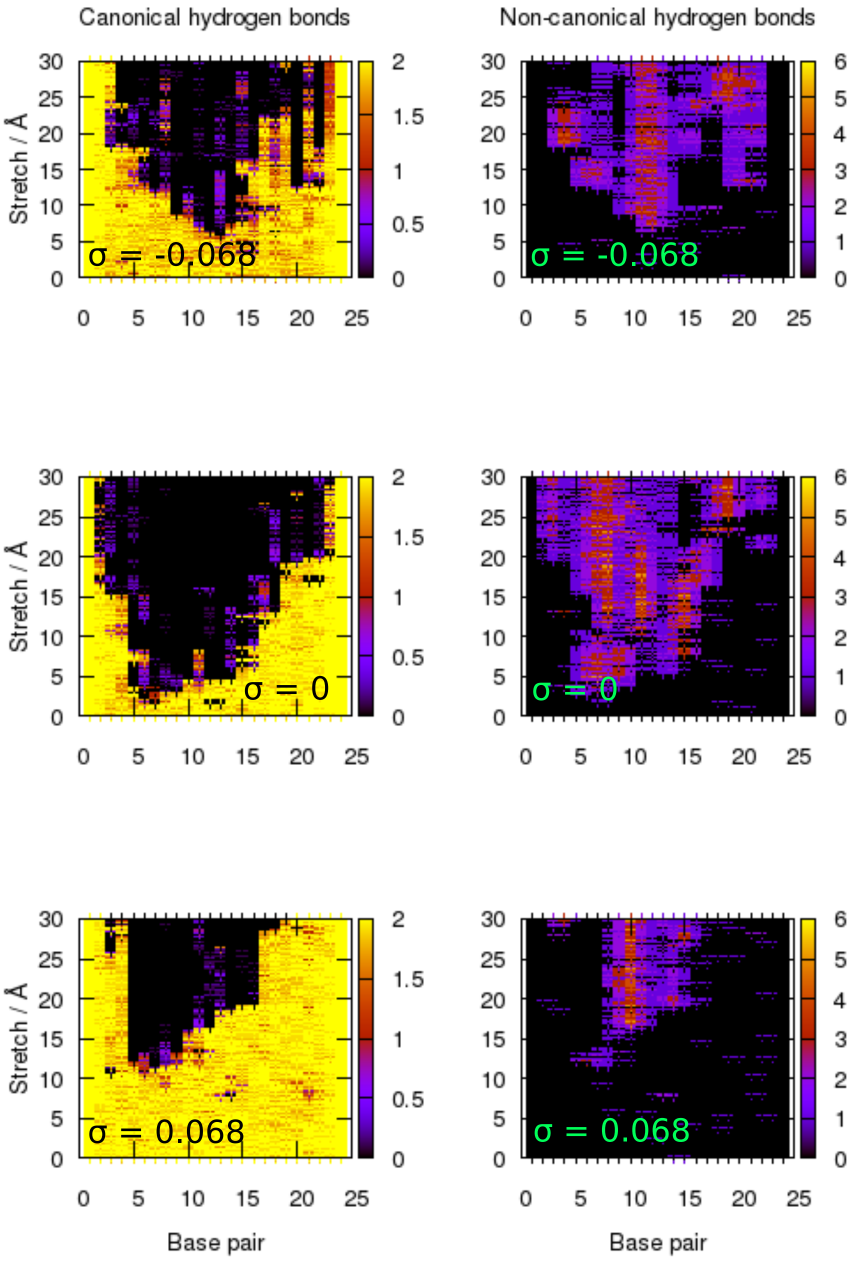


Figure B.3: poly-d(AT)·poly-d(AT) 100 ps/Å replica 2

## B.3 poly-d(C)·poly-d(G)

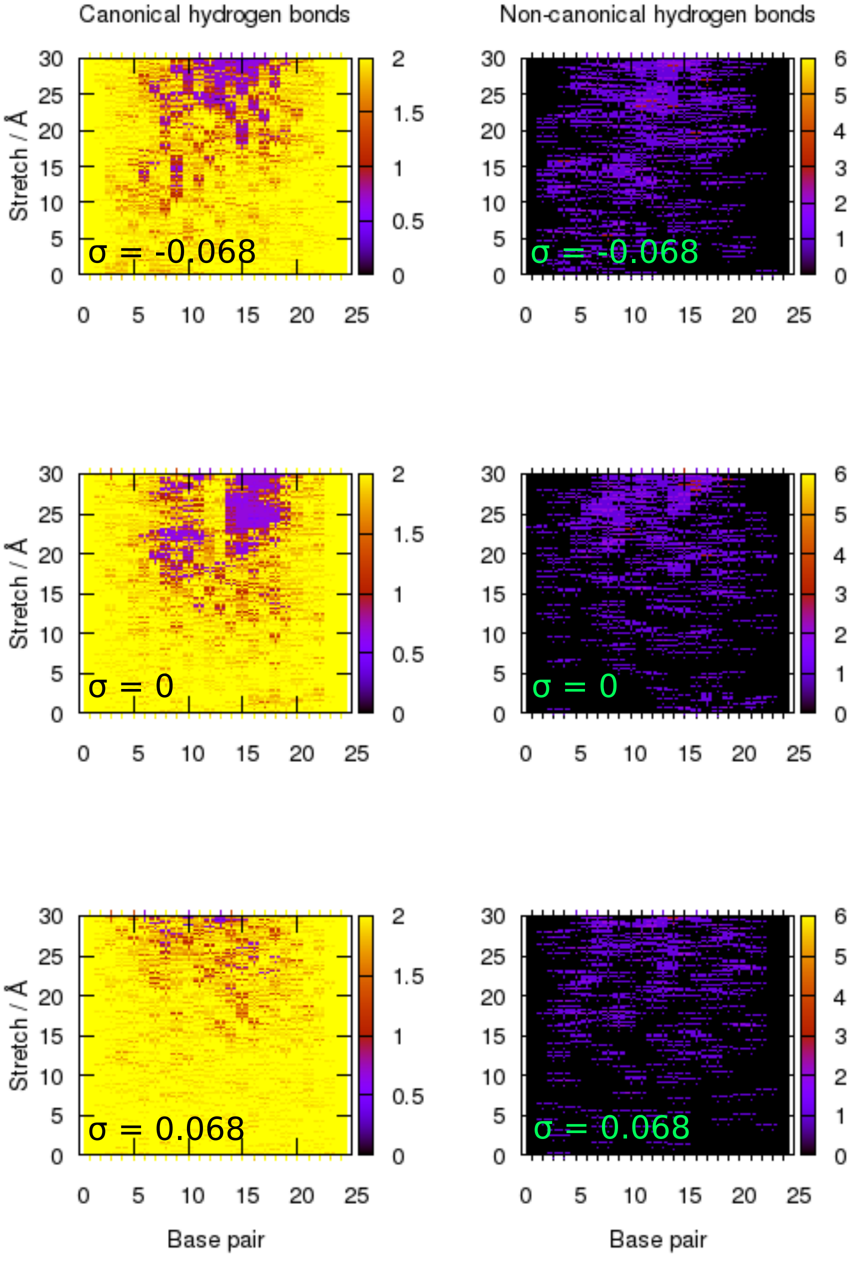


Figure B.4: poly-d(C)·poly-d(G) 100 ps/Å replica 1

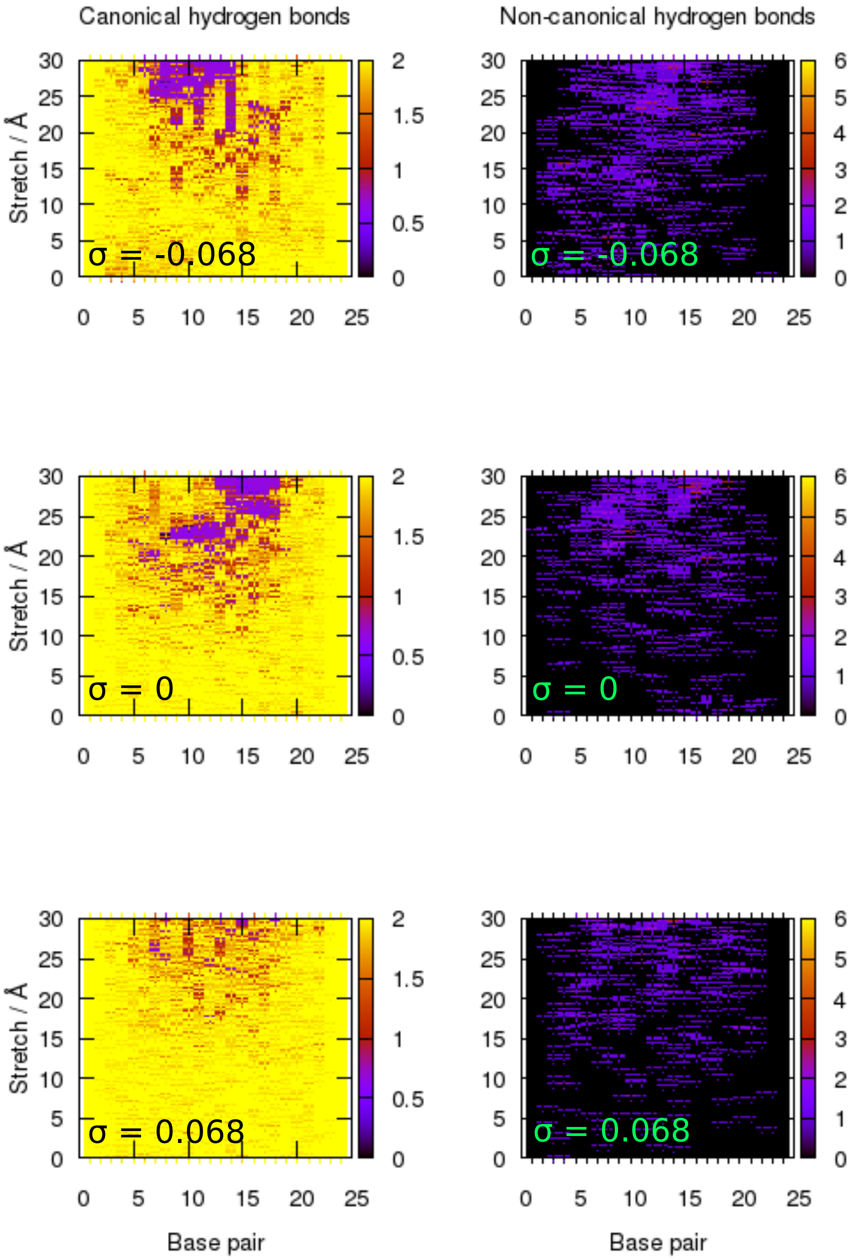


Figure B.5: poly-d(C)·poly-d(G) 100 ps/Å replica 2

## B.4 poly-d(CG)·poly-d(CG)

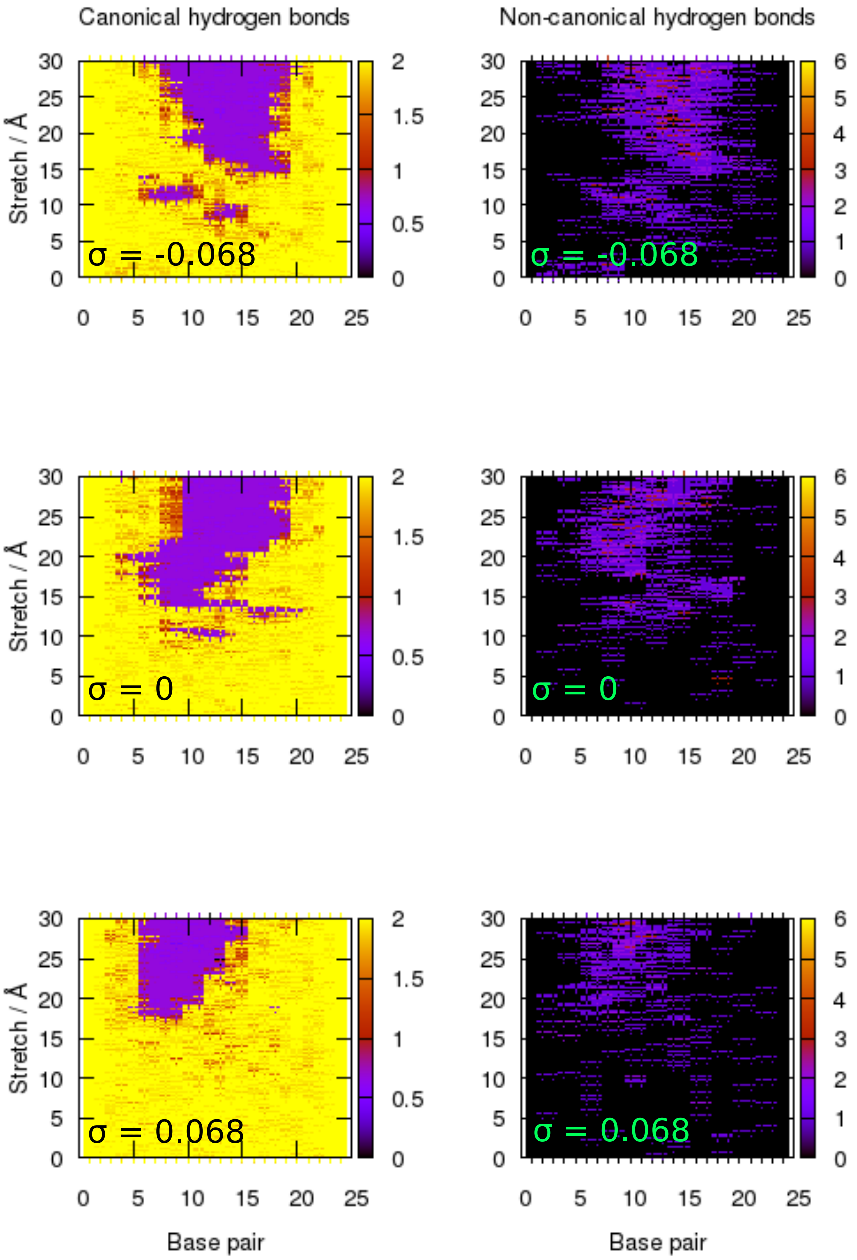


Figure B.6: poly-d(CG)·poly-d(CG) 100 ps/Å replica 1

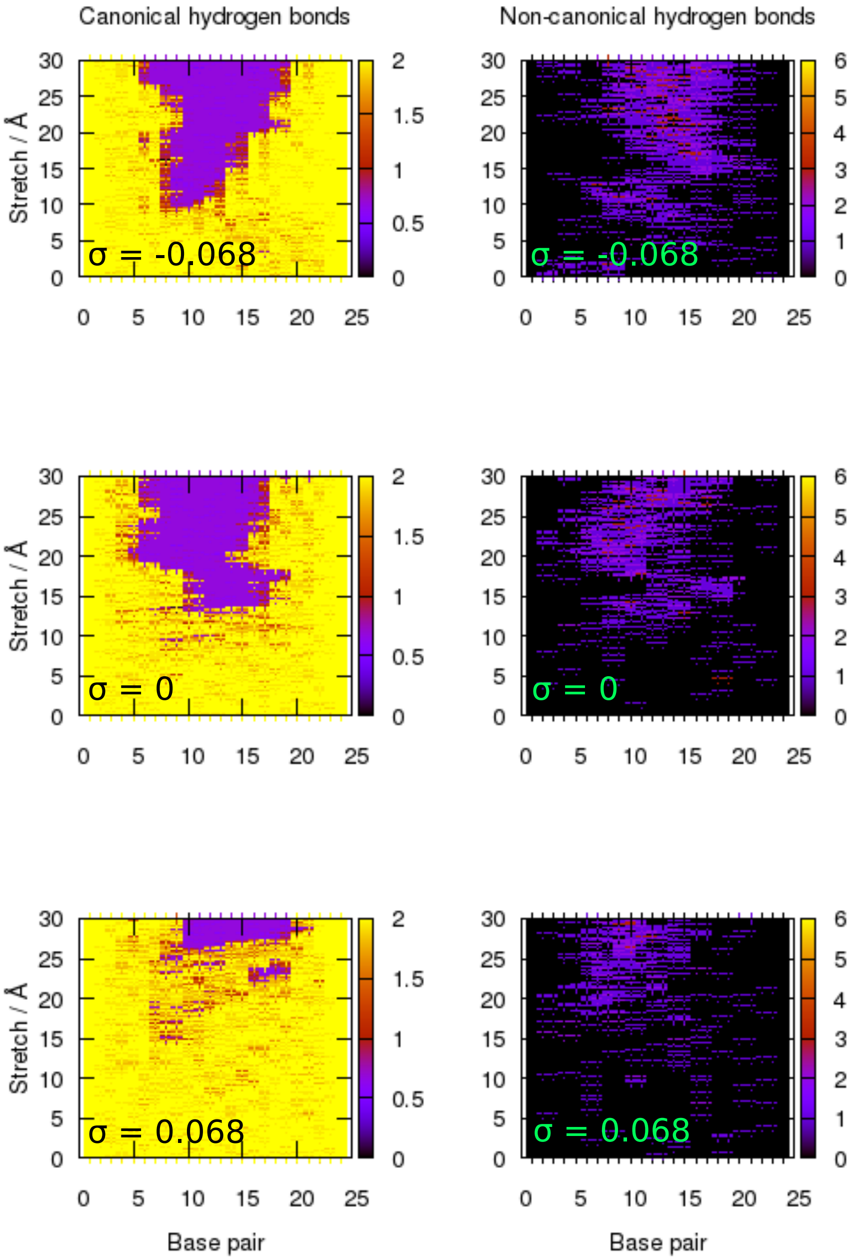


Figure B.7: poly-d(CG)·poly-d(CG) 100 ps/Å replica 2

# Bibliography

[1] Raimo Lohikoski, Jussi Timonen, and Aatto Laaksonen. Molecular dynamics simulation of single DNA stretching reveals a novel structure. *Chem. Phys. Lett.*, 407(1):23–29, 2005.

[2] Jean F Allemand, David Bensimon, Richard Lavery, and Vincent Croquette. Stretched and overwound DNA forms a Pauling-like structure with exposed bases. *Proc. Nat. Acad. Sci.*, 95(24):14152–14157, 1998.

[3] James D Watson and Francis HC Crick. The structure of DNA. In *Cold Spring Harbor Symposia on Quantitative Biology*, volume 18, pages 123–131. Cold Spring Harbor Laboratory Press, 1953.

[4] Carlos Bustamante, Steven B Smith, Jan Liphardt, and Doug Smith. Single-molecule studies of DNA mechanics. *Curr. Opin. Struct. Biol.*, 10(3):279–285, 2000.

[5] Scott Forth, Christopher Deufel, Maxim Y Sheinin, Bryan Daniels, James P Sethna, and Michelle D Wang. Abrupt buckling transition observed during the plectoneme formation of individual DNA molecules. *Phys. Rev. Lett.*, 100(14):148301, 2008.

[6] Bas Van Steensel and Job Dekker. Genomics tools for unraveling chromosome architecture. *Nat. Biotechnol.*, 28(10):1089, 2010.

[7] Christopher L Woodcock and Stefan Dimitrov. Higher-order structure of chromatin and chromosomes. *Curr. Opin. Genet. Dev.*, 11(2):130–135, 2001.

[8] Alexander Rich, Alfred Nordheim, and Andrew H-J Wang. The chemistry and biology of left-handed Z-DNA. *Annu. Rev. Biochem.*, 53(1):791–846, 1984.

[9] Alexander Rich and Shuguang Zhang. Z-DNA: the long road to biological function. *Nat. Rev. Genet.*, 4(7):566, 2003.

[10] Sung Chul Ha, Ky Lowenhaupt, Alexander Rich, Yang-Gyun Kim, and Kyeong Kyu Kim. Crystal structure of a junction between B-DNA and Z-DNA reveals two extruded bases. *Nature*, 437(7062):1183, 2005.

[11] Pui S Ho, Michael J Ellison, Gary J Quigley, and Alexander Rich. A computer aided thermodynamic approach for predicting the formation of Z-DNA in naturally occurring sequences. *EMBO J.*, 5(10):2737–2744, 1986.

[12] Graham L. Randall, Lynn Zechiedrich, and B. Montgomery Pettitt. In the absence of writhe, DNA relieves torsional stress with localized, sequence-dependent structural failure to preserve B-form. *Nucleic Acids Res.*, 37(16):5568–5577, 2009.

[13] G Eric Plum, Young-Whan Park, Scott F Singleton, Peter B Dervan, and Kenneth J Breslauer. Thermodynamic characterization of the stability and the melting behavior of a DNA triplex: a spectroscopic and calorimetric study. *Proc. Natl. Acad. Sci. U.S.A.*, 87(23):9436–9440, 1990.

[14] Anh Tuân Phan and Jean-Louis Mergny. Human telomeric DNA: G-quadruplex, i-motif and Watson–Crick double helix. *Nucleic Acids Res.*, 30(21):4618–4625, 2002.

[15] KJ Polach and J Widom. Mechanism of protein access to specific DNA sequences in chromatin: a dynamic equilibrium model for gene regulation. *J. Mol. Biol.*, 254(2):130–149, 1995.

[16] Keri Martinowich, Daisuke Hattori, Hao Wu, Shaun Fouse, Fei He, Yan Hu, Guoping Fan, and Yi E Sun. DNA methylation-related chromatin remodeling in activity-dependent BDNF gene regulation. *Science*, 302(5646):890–893, 2003.

[17] Gene-Wei Li, Otto G Berg, and Johan Elf. Effects of macromolecular crowding and dna looping on gene regulation kinetics. *Nat. Phys.*, 5(4):294, 2009.

[18] Craig Spiro, David P Bazett-Jones, Xiling Wu, and Cynthia T McMurray. DNA structure determines protein binding and transcriptional efficiency of the proenkephalin cAMP-responsive enhancer. *J. Biol. Chem.*, 270(46):27702–27710, 1995.

[19] William H Hoffman, Siham Biade, Jack T Zilfou, Jiandong Chen, and Maureen Murphy. Transcriptional repression of the anti-apoptoticsurvivin gene by wild type p53. *J. Biol. Chem.*, 277(5):3247–3257, 2002.

[20] Rolf Menzel and Martin Gellert. Regulation of the genes for E. coli DNA gyrase: homeostatic control of DNA supercoiling. *Cell*, 34(1):105–113, 1983.

[21] Steven P Gross, Michael A Welte, Steven M Block, and Eric F Wieschaus. Dynein-mediated cargo transport in vivo: a switch controls travel distance. *J. Cell Biol.*, 148(5):945–956, 2000.

[22] Leroy F Liu, Chung-Cheng Liu, and Bruce M Alberts. T4 DNA topoisomerase: a new ATP-dependent enzyme essential for initiation of T4 bacteriophage DNA replication. *Nature*, 281(5731):456, 1979.

[23] Connie Holm. Coming undone: how to untangle a chromosome. *Cell*, 77(7):955–957, 1994.

[24] James C Wang. Cellular roles of DNA topoisomerases: a molecular perspective. *Nat. Rev. Mol. Cell Biol.*, 3:430–440, 2002.

[25] Yves Pommier, Yilun Sun, Shar yin N. Huang, and John L. Nitiss. Roles of eukaryotic topoisomerases in transcription, replication and genomic stability. *Nat. Rev. Mol. Cell Biol.*, 17:703–721, 2016.

[26] Richard J Reece and Anthony Maxwell. DNA gyrase: structure and function. *Crit. Rev. Biochem. Mol. Biol.*, 26(3-4):335–375, 1991.

[27] John L Nitiss. Targeting DNA topoisomerase II in cancer chemotherapy. *Nat. Rev. Cancer*, 9(5):338, 2009.

[28] Yuliang Wu. Unwinding and Rewinding: Double Faces of Helicase? *J. Nucleic Acids*, 140601, 2012.

[29] Smita S. Patel1 and Ilker Donmez. Mechanisms of Helicases. *J. Biol. Chem.*, 281:18265–18268, 2006.

[30] D. J. Cole, C.-K. Skylaris, E. Rajendra, A. R. Venkitaraman, and M. C. Payne. Protein-protein interactions from linear-scaling first-principles quantum-mechanical calculations. *EPL*, 91, 2010.

[31] Thomas E. Ouldridge, Ard A. Louis, and Jonathan P. K. Doye. Structural, mechanical, and thermodynamic properties of a coarse-grained DNA model. *J. Chem. Phys*, 134, 2011.

[32] Thomas E. Ouldridge, Rollo L. Hoare, Ard A. Louis, Jonathan P. K. Doye, Jonathan Bath, and Andrew J. Turberfield. Optimizing DNA Nanotechnology through Coarse-Grained Modeling: A Two-Footed DNA Walker. *ACS Nano*, 7:2479–2490, 2013.

[33] AJ Spakowitz. Wormlike chain statistics with twist and fixed ends. *EPL*, 73(5):684, 2006.

[34] Kresten Lindorff-Larsen, Paul Maragakis, Stefano Piana, and David E. Shaw. Picosecond to Millisecond Structural Dynamics in Human Ubiquitin. *J. Phys. Chem. B*, 120:8313–8320, 2016.

[35] Levi C.T. Pierce, Romelia Salomon-Ferrer, Cesar Augusto F. de Oliveira, J. Andrew McCammon, and Ross C. Walker. Routine Access to Millisecond Time Scale Events with Accelerated Molecular Dynamics. *J. Chem. Theory Comput.*, 8:2997–3002, 2012.

[36] Kandadai N. Swamy and Enrico Clement. Hydration structure and dynamics of B- and Z-DNA in the presence of counterions via molecular dynamics simulations. *Biopolymers*, 26(11):1901–1927.

[37] W. F. Gunsteren, H. J. C. Berendsen, R. G. Geurtsen, and H. R. J. Zwinderman. A Molecular Dynamics Computer Simulation of an Eight-Base-Pair DNA Fragment in Aqueous Solution: Comparison with Experimental Two-Dimensional NMR Data. *Ann. N. Y. Acad. Sci.*, 482(1):287–303.

[38] Alan N. Real and Robert J. Greenall. Influence of Spermine on DNA Conformation in a Molecular Dynamics Trajectory of d(CGCGAATTCGCG)2: Major Groove Binding by One Spermine Molecule Delays the A-B Transition. *J. Biomol. Struct. Dyn.*, 21(4):469–487, 2004.

[39] Mats A. L. Eriksson, Torleif Hard, and Lennart Nilsson. Molecular Dynamics Simulations of the Glucocorticoid Receptor DNA-Binding Domain in Complex with DNA and Free in Solution. *Biophys. J.*, 68:402–426, 1995.

[40] Loïc Etheve, Juliette Martin, and Richard Lavery. Protein-DNA interfaces: a molecular dynamics analysis of time-dependent recognition processes for three transcription factors. *Nucleic Acids Res.*, 44(20):9990–10002, 2016.

[41] Alexander D MacKerell and Lennart Nilsson. Molecular dynamics simulations of nucleic acid-protein complexes. *Curr. Opin. Struc. Biol.*, 18(2):194–199, 2008.

[42] Helen Miller, Zhaokun Zhou, Jack Shepherd, Adam JM Wollman, and Mark C Leake. Single-molecule techniques in biophysics: a review of the progress in methods and applications. *Rep. Prog. Phys.*, 81(2):024601, 2017.

[43] Monika Fuxreiter, Mihaly Mezei, István Simon, and Roman Osman. Interfacial Water as a Hydration Fingerprint in the Noncognate Complex of BamHI. *Biophys. J.*, 89(2):903–911, 2005.

[44] Qing Zhang and Tamar Schlick. Stereochemistry and Position-Dependent Effects of Carcinogens on TATA/TBP Binding. *Biophys. J.*, 90(6):1865–1877, 2006.

[45] Niu Huang and Alexander D. MacKerell. Specificity in Protein-DNA Interactions: Energetic Recognition by the (Cytosine-C5)-methyltransferase from HhaI. *J. Mol. Biol.*, 345(2):265–274, 2005.

[46] Ramona Ettig, Nick Kepper, Rene Stehr, Gero Wedemann, and Karsten Rippe. Dissecting DNA-Histone Interactions in the Nucleosome by Molecular Dynamics Simulations of DNA Unwrapping. *Biophys. J.*, 101(8):1999–2008, 2011.

[47] Binquan Luan, Hongbo Peng, Stas Polonsky, Steve Rossnagel, Gustavo Stolovitzky, and Glenn Martyna. Base-By-Base Ratcheting of Single Stranded DNA through a Solid-State Nanopore. *Phys. Rev. Lett.*, 104:238103, 2010.

[48] M. Prabhakaran and Stephen C. Harvey. Molecular Dynamics Anneals Large-Scale Deformations of Model Macromolecules: Stretching the DNA Double Helix To Form an Intercalation Site. *J. Phys. Chem.*, 89:5767–5769, 1985.

[49] Alexander D. MacKerell Jr and Gil U. Lee. Structure, force, and energy of a double-stranded DNA oligonucleotide under tensile loads. *Eur. Biophys. J.*, 28:415–426, 1999.

[50] Michael W. Konrad and Joel I. Bolonick. Molecular Dynamics Simulation of DNA Stretching Is Consistent with the Tension Observed for Extension and Strand Separation and Predicts a Novel Ladder Structure. *J. Am. Chem. Soc.*, 118:10989–10994, 1996.

[51] Sarah A. Harris, Zara A. Sands, and Charles A. Laughton. Molecular Dynamics Simulations of Duplex Stretching Reveal the Importance of Entropy in Determining the Biomechanical Properties of DNA. *Biophys. J.*, 88(3):1684–1691, 2005.

[52] Alberto Marin-Gonzalez, J. G. Vilhena, Ruben Perez, and Fernando Moreno-Herrero. Understanding the mechanical response of double-stranded DNA and RNA under constant stretching forces using all-atom molecular dynamics. *Proc. Natl. Acad. Sci. U.S.A.*, 114(27):7049–7054, 2017.

[53] Srinivasaraghavan Kannan, Kai Kohlhoff, and Martin Zacharias. B-DNA Under Stress: Over- and Untwisting of DNA during Molecular Dynamics Simulations. *Biophys. J.*, 91(8):2956 – 2965, 2006.

[54] J. S. Mitchell, C. A. Laughton, and Sarah A. Harris. Atomistic simulations reveal bubbles, kinks and wrinkles in supercoiled DNA. *Nucleic Acids Res.*, 39(9):3928–3938, 2011.

[55] Andrew D. Bates, Agnes Noy, Michael M. Piperakis, Sarah A. Harris, and Anthony Maxwell. Small DNA circles as probes of DNA topology. *Biochem. Soc. Trans.*, 41:565–570, 2013.

[56] Jonathan M Fogg, Natalia Kolmakova, Ian Rees, Sergei Magonov, Helen Hansma, John J Perona, and E Lynn Zechiedrich. Exploring writhe in supercoiled minicircle DNA. *J. Phys Condens. Matter*, 18(14):S145, 2006.

[57] Jonathan S. Mitchell and Sarah A. Harris. Thermodynamics of Writhe in DNA Minicircles from Molecular Dynamics Simulations. *Phys. Rev. Lett.*, 110:148105, 2013.

[58] Andrew Travers and Georgi Muskhelishvili. A common topology for bacterial and eukaryotic transcription initiation? *EMBO Rep.*, 8(2):147–151, 2007.

[59] Thana Sutthibutpong, Christian Matek, Craig Benham, Gabriel G. Slade, Agnes Noy, Charles Laughton, Jonathan P. K. Doye, Ard A. Louis, and Sarah A. Harris. Long-range correlations in the mechanics of small DNA circles under topological stress revealed by multi-scale simulation. *Nucleic Acids Res.*, 44(19):9121–9130, 2016.

[60] Agnes Noy, Anthony Maxwell, and Sarah A. Harris. Interference between Triplex and Protein Binding to Distal Sites on Supercoiled DNA. *Biophys. J.*, 112(3):523–531, 2017.

[61] Thomas E. Ouldridge, Petr Šulc, Flavio Romano, Jonathan P. K. Doye, and Ard A. Louis. DNA hybridization kinetics: zippering, internal displacement and sequence dependence. *Nucleic Acids Res.*, 41(19):8886–8895, 2013.

[62] Majid Mosayebi, Ard A. Louis, Jonathan P. K. Doye, and Thomas E. Ouldridge. Force-Induced Rupture of a DNA Duplex: From Fundamentals to Force Sensors. *ACS Nano*, 9(12):11993–12003, 2015.

[63] R. M. Harrison, F. Romano, T. E. Ouldridge, A. A. Louis, and J. P. K. Doye. Coarse-grained modelling of strong DNA bending I: Thermodynamics and comparison to an experimental “molecular vice”. *ArXiv e-prints*, 2015.

[64] R. M. Harrison, F. Romano, T. E. Ouldridge, A. A. Louis, and J. P. K. Doye. Coarse-grained modelling of strong DNA bending II: Cyclization. *ArXiv e-prints*, 2015.

[65] Christian Matek, Thomas E. Ouldridge, Jonathan P. K. Doye, and Ard A. Louis. Plectoneme tip bubbles: Coupled denaturation and writhing in supercoiled DNA. *Sci. Rep.*, 5:7655, 2015.

[66] Flavio Romano, Debayan Chakraborty, Jonathan P. K. Doye, Thomas E. Ouldridge, and Ard A. Louis. Coarse-grained simulations of DNA overstretching. *J. Chem. Phys.*, 138(8):085101, 2013.

[67] Petr Šulc, Flavio Romano, Thomas E. Ouldridge, Lorenzo Rovigatti, Jonathan P. K. Doye, and Ard A. Louis. Sequence-dependent thermodynamics of a coarse-grained DNA model. *J. Chem. Phys.*, 137(13):135101, 2012.

[68] Benedict E. K. Snodin, Ferdinando Randisi, Majid Mosayebi, Petr Šulc, John S. Schreck, Flavio Romano, Thomas E. Ouldridge, Roman Tsukanov, Eyal Nir, Ard A. Louis, and Jonathan P. K. Doye. Introducing improved structural properties and salt dependence into a coarse-grained model of DNA. *J. Chem. Phys.*, 142(23):234901, 2015.

[69] Enrico Skoruppa, Michiel Laleman, Stefanos K. Nomidis, and Enrico Carlon. DNA elasticity from coarse-grained simulations: The effect of groove asymmetry. *J. Chem. Phys.*, 146(21):214902, 2017.

[70] Christoph G Baumann, Steven B Smith, Victor A Bloomfield, and Carlos Bustamante. Ionic effects on the elasticity of single DNA molecules. *Proc. Natl. Acad. Sci. U.S.A.*, 94(12):6185–6190, 1997.

[71] Qian Wang and B. Montgomery Pettitt. Modeling DNA Thermodynamics under Torsional Stress. *Biophys. J.*, 106(5):1182–1193, 2014.

[72] E. Skoruppa, S. K. Nomidis, J. F. Marko, and E. Carlon. Bend-Induced Twist Waves and the Structure of Nucleosomal DNA. *ArXiv e-prints*, 2018.

[73] John F Marko and Eric D Siggia. Stretching DNA. *Macromolecules*, 28(26):8759–8770, 1995.

[74] Rafayel Petrosyan. Improved approximations for some polymer extension models. *Rheol. Acta*, 56(1):21–26, 2017.

[75] Theo Odijk. Stiff chains and filaments under tension. *Macromolecules*, 28(20):7016–7018, 1995.

[76] Peter Gross, Niels Laurens, Lene B Oddershede, Ulrich Bockelmann, Erwin JG Peterman, and Gijs JL Wuite. Quantifying how DNA stretches, melts and changes twist under tension. *Nat. Phys.*, 7(9):731, 2011.

[77] Jens-Christian Meiners and Stephen R. Quake. Femtonewton Force Spectroscopy of Single Extended DNA Molecules. *Phys. Rev. Lett.*, 84:5014–5017, May 2000.

[78] Frederick Gittes and Christoph F. Schmidt. Interference model for back-focal-plane displacement detection in optical tweezers. *Opt. Lett.*, 23(1):7–9, Jan 1998.

[79] Matthew J Lang, Charles L Asbury, Joshua W Shaevitz, and Steven M Block. An automated two-dimensional optical force clamp for single molecule studies. *Biophys. J.*, 83(1):491–501, 2002.

[80] SB Smith, L Finzi, and C Bustamante. Direct mechanical measurements of the elasticity of single DNA molecules by using magnetic beads. *Science*, 258(5085):1122–1126, 1992.

[81] M.D. Wang, H. Yin, R. Landick, J. Gelles, and S.M. Block. Stretching DNA with optical tweezers. *Biophys. J.*, 72(3):1335 – 1346, 1997.

[82] U. Bockelmann, Ph. Thomen, B. Essevaz-Roulet, V. Viasnoff, and F. Heslot. Unzipping DNA with Optical Tweezers: High Sequence Sensitivity and Force Flips. *Biophys. J.*, 82(3):1537 – 1553, 2002.

[83] Martin L. Bennink, Orlando D. Scharer, Roland Kanaar, Kumiko Sakata-Sogawa, Juleon M. Schins, Johannes S. Kanger, Bart G. de Grooth, and Jan Greve. Single-molecule manipulation of double-stranded DNA using optical tweezers: Interaction studies of DNA with RecA and YOYO-1. *Cytometry*, 36(3):200–208.

[84] Katrin Günther, Michael Mertig, and Ralf Seidel. Mechanical and structural properties of YOYO-1 complexed DNA. *Nucleic Acids Res.*, 38(19):6526–6532, 2010.

[85] Binu Kundukad, Jie Yan, and Patrick S Doyle. Effect of YOYO-1 on the mechanical properties of DNA. *Soft Matter*, 10(48):9721–9728, 2014.

[86] C.U. Murade, V. Subramaniam, C. Otto, and Martin L. Bennink. Interaction of Oxazole Yellow Dyes with DNA Studied with Hybrid Optical Tweezers and Fluorescence Microscopy. *Biophys. J.*, 97(3):835 – 843, 2009.

[87] Xiaoyan R. Bao, Heun Jin Lee, and Stephen R. Quake. Behavior of Complex Knots in Single DNA Molecules. *Phys. Rev. Lett.*, 91:265506, Dec 2003.

[88] Shaevitz, Joshua W. and Abbondanzieri, Elio A. and Landick, Robert and Block, Steven M. Backtracking by single RNA polymerase molecules observed at near-base-pair resolution. *Nature*, 426:684–687, 2003.

[89] Elio A. Abbondanzieri, William J. Greenleaf, Joshua W. Shaevitz, Robert Landick, and Steven M. Block. Direct observation of base-pair stepping by RNA polymerase. *Nature*, 438:460–465, 2005.

[90] Chih-Ming Cheng, Yuarn-Jang Lee, Wei-Ting Wang, Chien-Ting Hsu, Jing-Shin Tsai, Chien-Ming Wu, Keng-Liang Ou, and Tzu-Sen Yang. Determining the binding mode and binding affinity constant of tyrosine kinase inhibitor PD153035 to DNA using optical tweezers. *Biochem. Biophy. Res. Commun.*, 404(1):297 – 301, 2011.

[91] Iddo Heller, Gerrit Sitters, Onno D Broekmans, Géraldine Farge, Carolin Menges, Wolfgang Wende, Stefan W Hell, Erwin J G Peterman, and Gijs J L Wuite. STED nanoscopy combined with optical tweezers reveals protein dynamics on densely covered DNA. *Nat. Methods*, 10:910–916, 2013.

[92] Terence R Strick, J-F Allemand, David Bensimon, Aaron Bensimon, and Vincent Croquette. The elasticity of a single supercoiled DNA molecule. *Science*, 271(5257):1835–1837, 1996.

[93] JF Leger, G Romano, A Sarkar, J Robert, L Bourdieu, D Chatenay, and JF Marko. Stru ctural transitions of a twisted and stretched DNA molecule. *Phys. Rev. Lett.*, 83(5):1066, 1999.

[94] Terence R Strick, Vincent Croquette, and David Bensimon. Single-molecule analysis of DNA uncoiling by a type II topoisomerase. *Nature*, 404(6780):901, 2000.

[95] G Charvin, TR Strick, D Bensimon, and V Croquette. Topoisomerase IV bends and overtwists DNA upon binding. *Biophys. J.*, 89(1):384–392, 2005.

[96] Jeff Gore, Zev Bryant, Marcelo Nöllmann, Mai U Le, Nicholas R Cozzarelli, and Carlos Bustamante. DNA overwinds when stretched. *Nature*, 442(7104):836, 2006.

[97] Paul Lebel, Aakash Basu, Florian C Oberstrass, Elsa M Tretter, and Zev Bryant. Gold rotor bead tracking for high-speed measurements of DNA twist, torque and extension. *Nat. Methods*, 11(4):456, 2014.

[98] Jan Lipfert, Matthew Wiggin, Jacob WJ Kerssemakers, Francesco Pedaci, and Nynke H Dekker. Freely orbiting magnetic tweezers to directly monitor changes in the twist of nucleic acids. *Nat. Commun.*, 2:439, 2011.

[99] Maxim Y Sheinin and Michelle D Wang. Twist–stretch coupling and phase transition during DNA supercoiling. *Phys. Chem. Chem. Phys.*, 11(24):4800–4803, 2009.

[100] Domenico Salerno, Doriano Brogioli, Valeria Cassina, Diana Turchi, Giovanni Luca Beretta, Davide Seruggia, Roberto Ziano, Franco Zunino, and Francesco Mantegazza. Magnetic tweezers measurements of the nanomechanical properties of DNA in the presence of drugs. *Nucleic Acids Res.*, 38(20):7089–7099, 2010.

[101] Jan Lipfert, Jacob WJ Kerssemakers, Tessa Jager, and Nynke H Dekker. Magnetic torque tweezers: measuring torsional stiffness in DNA and RecA-DNA filaments. *Nat. Methods*, 7(12):977, 2010.

[102] Sébastien Neukirch. Extracting DNA twist rigidity from experimental supercoiling data. *Phys. Rev. Lett.*, 93(19):198107, 2004.

[103] Francesco Mosconi, Jean François Allemand, David Bensimon, and Vincent Croquette. Measurement of the torque on a single stretched and twisted DNA using magnetic tweezers. *Phys. Rev. Lett.*, 102(7):078301, 2009.

[104] Thomas T Perkins, DE Smith, S Chu, et al. Relaxation of a single DNA molecule observed by optical microscopy. *Science*, 264(5160):822–826, 1994.

[105] D Bensimon, AJ Simon, V Croquette, and A Bensimon. Stretching DNA with a receding meniscus: experiments and models. *Phys. Rev. Lett.*, 74(23):4754, 1995.

[106] Jiunn B Heng, Aleksei Aksimentiev, Chuen Ho, Patrick Marks, Yelena V Grinkova, Steve Sligar, Klaus Schulten, and Gregory Timp. Stretching DNA using the electric field in a synthetic nanopore. *Nano Lett.*, 5(10):1883–1888, 2005.

[107] Andrew Hards, Chunqing Zhou, Markus Seitz, Christoph Bräuchle, and Andreas Zumbusch. Simultaneous AFM manipulation and fluorescence imaging of single DNA strands. *ChemPhysChem*, 6(3):534–540, 2005.

[108] Hiroki Yokota, James Sunwoo, Mehmet Sarikaya, Ger van den Engh, and Ruedi Aebersold. Spin-stretching of DNA and protein molecules for detection by fluorescence and atomic force microscopy. *Anal. Chem.*, 71(19):4418–4422, 1999.

[109] Zhaokun Zhou, Helen Miller, Adam JM Wollman, and Mark C Leake. Developing a new biophysical tool to combine magneto-optical tweezers with super-resolution fluorescence microscopy. In *Photonics*, volume 2, pages 758–772. Multidisciplinary Digital Publishing Institute, 2015.

[110] Yeonee Seol and Keir C Neuman. Single-molecule measurements of topoisomerase activity with magnetic tweezers. In *Single Molecule Enzymology: Methods and Protocols*, chapter 15, pages 229–241. Springer, 2011.

[111] Zhaokun Zhou. *Magneto-optical tweezers with super-resolution fluorescence microscopy.* PhD Thesis, University of York, 2017.

[112] Ignacio J General, Ralitsa Dragomirova, and Hagai Meirovitch. Absolute free energy of binding of avidin/biotin, revisited. *J. Phys. Chem. B*, 116(23):6628–6636, 2012.

[113] ThermoFisher DNA and RNA Molecular Weights and Conversions Tables. `https://www.thermofisher.com/uk/en/home/references/ambion-tech-support/rna-tools-and-calculators/dna-and-rna-molecular-weights-and-conversions.html`. Retrieved 28/07/2018.

[114] NEB Lambda DNA product page. `https://international.neb.com/products/n3011-lambda-dna#Product%20Information`. Retrieved 28/07/2018.

[115] Elio A Abbondanzieri, William J Greenleaf, Joshua W Shaevitz, Robert Landick, and Steven M Block. Direct observation of base-pair stepping by RNA polymerase. *Nature*, 438(7067):460, 2005.

[116] Mark Bates, Sara A Jones, and Xiaowei Zhuang. Stochastic optical reconstruction microscopy (STORM): a method for superresolution fluorescence imaging. *Cold Spring Harb. Protoc.*, 2013(6):pdb–top075143, 2013.

[117] ThermoFisher Scientific YOYO-1 product page. `https://www.thermofisher.com/order/catalog/product/Y3601`. Retrieved 27/08/2018.

[118] Semrock BrightLine product page. `https://www.semrock.com/setdetails.aspx?id=2711`. Retrieved 27/08/2018.

[119] Helen Miller, Zhaokun Zhou, Adam J.M. Wollman, and Mark C. Leake. Superresolution imaging of single DNA molecules using stochastic photoblinking of minor groove and intercalating dyes. *Methods*, 88:81–88, 2015.

[120] Hays S Rye, Stephen Yue, David E Wemmer, Mark A Quesada, Richard P Haugland, Richard A Mathies, and Alexander N Glazer. Stable fluorescent complexes of double-stranded DNA with bis-intercalating asymmetric cyanine dyes: properties and applications. *Nucleic Acids Res.*, 20(11):2803–2812, 1992.

[121] Sebastian van de Linde and Markus Sauer. How to switch a fluorophore: from undesired blinking to controlled photoswitching. *Chem. Soc. Rev.*, 43(4):1076–1087, 2014.

[122] Thorben Cordes, Jan Vogelsang, and Philip Tinnefeld. On the mechanism of Trolox as antiblinking and antibleaching reagent. *J. Am. Chem. Soc.*, 131(14):5018–5019, 2009.

[123] Andrea Candelli, Gijs JL Wuite, and Erwin JG Peterman. Combining optical trapping, fluorescence microscopy and micro-fluidics for single molecule studies of DNA–protein interactions. *Phys. Chem. Chem. Phys.*, 13(16):7263–7272, 2011.

[124] Specifications of T4 DNA Ligase. `https://international.neb.com/products/m0202-t4-dna-ligase#Product%20Information_Properties%20and%20Usage`. Retrieved 12/06/2018.

[125] Components of NEB Ligase Buffer. `https://www.neb.com/products/b0202-t4-dna-ligase-reaction-buffer`. Retrieved 12/06/2018.

[126] Caroline A Schneider, Wayne S Rasband, and Kevin W Eliceiri. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods*, 9(7):671, 2012.

[127] Andor Solis product page. `https://andor.oxinst.com/products/solis-software/solis-i`. Retrieved 27/08/2018.

[128] Mad City Labs Nano-LP range product page. `http://www.madcitylabs.com/nanolpseries.html`. Retrieved 27/08/2018.

[129] Keir C Neuman and Steven M Block. Optical trapping. *Rev. Sci. Instrum.*, 75(9):2787–2809, 2004.

[130] Thomas T Perkins. Optical traps for single molecule biophysics: a primer. *Laser Photonics Rev.*, 3(1-2):203–220, 2009.

[131] Jeffrey R Moffitt, Yann R Chemla, Steven B Smith, and Carlos Bustamante. Recent advances in optical tweezers. *Annu. Rev. Biochem.*, 77:205–228, 2008.

[132] A Mazolli, PA Maia Neto, and HM Nussenzveig. Theory of trapping forces in optical tweezers. In *Proc. Royal. Soc. Lond. A*, volume 459, pages 3021–3041. The Royal Society, 2003.

[133] Mohammed Mahamdeh, Citlali Pérez Campos, and Erik Schäffer. Under-filling trapping objectives optimizes the use of the available laser power in optical tweezers. *Optics express*, 19(12):11759–11768, 2011.

[134] Thorlabs QPD product page. `https://www.thorlabs.com/newgrouppage9.cfm?objectgroup_id=4400`. Retrieved 27/08/2018.

[135] Kirstine Berg-Sørensen and Henrik Flyvbjerg. Power spectrum analysis for optical tweezers. *Rev. Sci. Instrum.*, 75(3):594–612, 2004.

[136] MIM Feitosa and ON Mesquita. Wall-drag effect on diffusion of colloidal particles near surfaces: a photon correlation study. *Phys. Rev. A*, 44(10):6677, 1991.

[137] Yeonee Seol, Amanda E Carpenter, and Thomas T Perkins. Gold nanoparticles: enhanced optical trapping and sensitivity coupled with significant heating. *Opt. Lett.*, 31(16):2429–2431, 2006.

[138] SpheroTech Customer Services. Private Communication.

[139] Stephen John Cross. *Combining magnetic tweezers and single-molecule fluorescence microscopy to probe transcription-coupled DNA supercoiling.* PhD Thesis, University of York, 2013.

[140] CJ Van Oss, RJ Good, and MK Chaudhury. Mechanism of DNA (Southern) and protein (Western) blotting on cellulose nitrate and other membranes. *J. Chromatogr. A*, 391:53–65, 1987.

[141] ThermoFisher PBS Formulation. `http://www.thermofisher.com/uk/en/home/technical-resources/media-formulation.317.html`. Retrieved 30/07/2018.

[142] Erwin JG Peterman, Frederick Gittes, and Christoph F Schmidt. Laser-induced heating in optical traps. *Biophys. J.*, 84(2):1308–1316, 2003.

[143] Annaël Brunet, Catherine Tardin, Laurence Salome, Philippe Rousseau, Nicolas Destainville, and Manoel Manghi. Dependence of DNA persistence length on ionic strength of solutions with monovalent and divalent salts: a joint theory–experiment study. *Macromolecules*, 48(11):3641–3652, 2015.

[144] TR Strick, J-F Allemand, D Bensimon, and V Croquette. Behavior of supercoiled DNA. *Biophys. J.*, 74(4):2016–2028, 1998.

[145] Rifka Vlijm, Alireza Mashaghi, Stéphanie Bernard, Mauro Modesti, and Cees Dekker. Experimental phase diagram of negatively supercoiled DNA measured by magnetic tweezers and fluorescence. *Nanoscale*, 7(7):3205–3216, 2015.

[146] D.A. Case, R.M. Betz, D.S. Cerutti, T.E. Cheatham III, T.A. Darden, R.E. Duke, T.J. Giese, H. Gohlke, A.W. Goetz, N. Homeyer, S. Izadi, P. Janowski, J. Kaus, A. Kovalenko, T.S. Lee, S. LeGrand, P. Li, C. Lin, T. Luchko, R. Luo, B. Madej, D. Mermelstein, K.M. Merz, G. Monard, H. Nguyen, H.T. Nguyen, I. Omelyan, A. Onufriev, D.R. Roe, A. Roitberg, C. Sagui, C.L. Simmerling, W.M. Botello-Smith, J. Swails, R.C. Walker, J. Wang, R.M. Wolf, X. Wu, L. Xiao, and P.A. Kollman. *Amber 16.* University of California, San Francisco, 2016.

[147] John Edward Lennard-Jones. On the determination of molecular fields. II. From the equation of state of gas. *Proc. Roy. Soc. A*, 106:463–477, 1924.

[148] Tom Darden, Darrin York, and Lee Pedersen. Particle mesh Ewald: An Nlog(N) method for Ewald sums in large systems. *J. Chem. Phys.*, 98(12):10089–10092, 1993.

[149] Ivan Ivani, Pablo D Dans, Agnes Noy, Alberto Pérez, Ignacio Faustino, Adam Hospital, Jürgen Walther, Pau Andrio, Ramon Goñi, Alexandra Balaceanu, et al. Parmbsc1: a refined force field for DNA simulations. *Nat. Methods*, 13(1):55, 2016.

[150] Alberto Pérez, Iván Marchán, Daniel Svozil, Jiri Sponer, Thomas E Cheatham III, Charles A Laughton, and Modesto Orozco. Refinement of the AMBER force field for nucleic acids: improving the description of $\alpha/\gamma$ conformers. *Biophys. J.*, 92(11):3817–3829, 2007.

[151] Junmei Wang, Piotr Cieplak, and Peter A Kollman. How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J. Comput. Chem.*, 21(12):1049–1074, 2000.

[152] Rodrigo Galindo-Murillo, James C Robertson, Marie Zgarbova, Jirí Sponer, Michal Otyepka, Petr Jurecka, and Thomas E Cheatham III. Assessing the current state of amber force field modifications for DNA. *J. Chem. Theory Comput.*, 12(8):4114–4127, 2016.

[153] Interactions in the oxDNA model. `https://dna.physics.ox.ac.uk/index.php/DNA_model_introduction`. Retrieved 12/06/18.

[154] Thomas E. Ouldridge. *Coarse-grained modelling of DNA and DNA self-assembly.* DPhil Thesis, University of Oxford, 2011.

[155] R. J. Loncharich, B. R. Brooks, and R. W. Pastor. Langevin dynamics of peptides: The frictional dependence of isomerization rates of n-acetylalnyl-n'-methylamide. *Biopolymers*, 32:523–535, 1992.

[156] W Clark Still, Anna Tempczyk, Ronald C Hawley, and Thomas Hendrickson. Semianalytical treatment of solvation for molecular mechanics and dynamics. *J. Am. Chem. Soc.*, 112(16):6127–6129, 1990.

[157] Hai Nguyen, Alberto Peerez, Sherry Bermeo, and Carlos Simmerling. Refinement of generalized born implicit solvation parameters for nucleic acids and their complexes with proteins. *J. Chem. Theory Comput.*, 11(8):3714–3728, 2015.

[158] Alexey Onufriev, David A Case, and Donald Bashford. Effective Born radii in the generalized Born approximation: the importance of being perfect. *J. Comput. Chem.*, 23(14):1297–1304, 2002.

[159] Jayashree Srinivasan, Megan W Trevathan, Paul Beroza, and David A Case. Application of a pairwise generalized Born model to proteins and nucleic acids: inclusion of salt effects. *Theor. Chem. Acc.*, 101(6):426–434, 1999.

[160] Alexander Stukowski. Visualization and analysis of atomistic simulation data with OVITO–the Open Visualization Tool. *Model. Simul. Mater. Sci. Eng.*, 18(1):015012, 2009.

[161] Ovito Homepage. `http://ovito.org`. Retrieved 8/08/2018.

[162] William Humphrey, Andrew Dalke, and Klaus Schulten. VMD – Visual Molecular Dynamics. *J. Mol. Graph.*, 14:33–38, 1996.

[163] Sean A McKinney, Chirlmin Joo, and Taekjip Ha. Analysis of single-molecule FRET trajectories using hidden Markov modeling. *Biophys. J.*, 91(5):1941–1951, 2006.

[164] Julius Marmur and Paul Doty. Determination of the base composition of deoxyribonucleic acid from its thermal denaturation temperature. *J. Mol. Biol.*, 5(1):109–118, 1962.

[165] Daniel R Roe and Thomas E Cheatham III. PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data. *J. Chem. Theory Comput.*, 9(7):3084–3095, 2013.

[166] Amit D Mehta, Matthias Rief, James A Spudich, David A Smith, and Robert M Simmons. Single-molecule biomechanics with optical methods. *Science*, 283(5408):1689–1695, 1999.