Eindhoven University of Technology

MASTER

DNA catch bonds

theory, simulation, and experiments on a simple, humanmade, and tunable catch bond

van Leuken, Stijn

Award date:
2019

Link to publication

Disclaimer
This document contains a student thesis (bachelor's or master's), as authored by a student at Eindhoven University of Technology. Student theses are made available in the TU/e repository upon obtaining the required degree. The grade received is not published on the document as presented in the repository. The required complexity or quality of research of student theses may vary by program, and the required minimum study period may vary in duration.

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain

Eindhoven University of Technology

DNA Catch Bonds

Theory, Simulation, and Experiments on a Simple, Humanmade, and Tunable Catch Bond

Master Thesis
by:

Stijn H. M. van Leuken

January 17, 2019

Abstract

Ordinary non-covalent bonds unbind faster as the force on them is increased. These so-called slip bonds occur throughout biology, chemistry, and physics. For catch bonds the inverse is true: their lifetime grows with applied tension. Experiments report this counterintuitive property in a wide range of biological receptor-ligand pairs. A number of underlying physical scenarios for this unexpected behavior have been put forward. However, the complexity of the involved molecules generally makes it difficult to identify which of these scenarios occurs in a particular system. What is clear, is that catch bonds provide nonstandard mechanical functionality at molecular scales. Mimicking catch bond behavior in well-controlled humanmade constructs makes such functionality available in synthetic materials. In our work, we have assessed the possibility to create a catch bond using DNA. The simplicity and tunability of our construct allow confronting theory with observations as well as permitting material design with unique properties. In this master thesis, I combine theory, experiment, and simulation to examine whether the designed DNA construct can be used to create predictable and tunable catch bonds. Our theory, numerical simulations and experiments all show signatures of catch bonding, suggesting it is possible to both understand and engineer catch bond functionality with DNA.

Contents

1 Introduction: Catch bonds and DNA

Non-covalent can be classified into two groups: slip bonds and catch bonds. These bonds respond differently to forces. The higher the force applied to a slip bond, the faster it dissociates. In contrast, the lifetime of catch bonds increases with increased tension. This section introduces catch bonds by discussing experimental observations and possible physical explanations of these observations.

Replicating the nonstandard lifetime-force dependency observed in biological experiments in simple, humanmade constructs makes it possible to use catch bonds in synthetic materials. In this M.Sc. thesis, I ask the question whether a predictable and tunable catch bond can be created using DNA. In the second half of the introduction, I explain why DNA is a suitable molecule for this purpose and how existing models can be used to predict its response to force.

1.1 Observations, physical scenarios of Catch bonds

Experimental observations of catch bonds Catch bonding has been reported in a variety of biological systems. B. Savage et al., B.J. Fredrickson et al. and M.H. Kroll et al. observed a catch bond behavior between platelet and a surface coated with plasma protein [2–4]. A second example is the bond between leukocytes (P-Selectin) and P-Selectin Glycoprotein Ligand-1 (PSGL) [1, 5]. Also, the agglutination of red blood cells shows signatures of catch bond behavior [6]. Z.J. Li et al. showed that the bond between bacterium Staphylococcus Aureus and a collagen coated surface behaves as a catch bond [7]. Also, the bond between type 1 Fimbrin D and Mannose in Escherichia Coli [8], the bond between Fibronectin and an Integrin $\alpha_5\beta_1 - Fc$ fusion protein or membrane [9] and the bond between Vinculin and F-Actin [10] have an inversed lifetime-force dependency. As a final example, Catherin cell-cell adhesion shows signatures of catch bonding [11]. This list is not exhaustive and more catch bonds in biological systems could be observed in the future.

As an example, I discuss the bond between P-Selectin and PSGL. Leukocytes move through the bloodstream as a result of the flow of blood. When one of these leukocytes hits the vessel walls, a connection between the leukocyte and the vessel wall can be formed via a P-Selectin-PSGL bond. In some cases, such a connection breaks and the leukocyte continues to move in the bloodstream. In other cases, multiple connections between the leukocyte and the blood vessel form, and by breakage and reformation of connections at slightly different locations, the leukocyte moves along the surface in a rotating motion called leukocyte rolling. The higher the speed of flow in the bloodstream, the higher the force pulling on the attachment between the leukocytes and the vessel wall. Counter-intuitively, for a range of flow rates, the average time spent by a leukocyte attached to a wall increases with the applied force.

The observation of rolling leukocytes suggests that the P-Selectin-PSGL bond attaching leukocytes to vessel walls is a catch bond. Marshall et al. measured the lifetime of this bond as a function of force experimentally using atomic force microscopy. Figure 1 shows their results. With these experiments, Marshall et al. showed that there is a range of forces for the P-Selectin-PSGL bond where lifetime increases with force. So their results suggest the bond is a catch bond. This result, together with other measured catch bonds effects shows that catch bonds exist in nature.

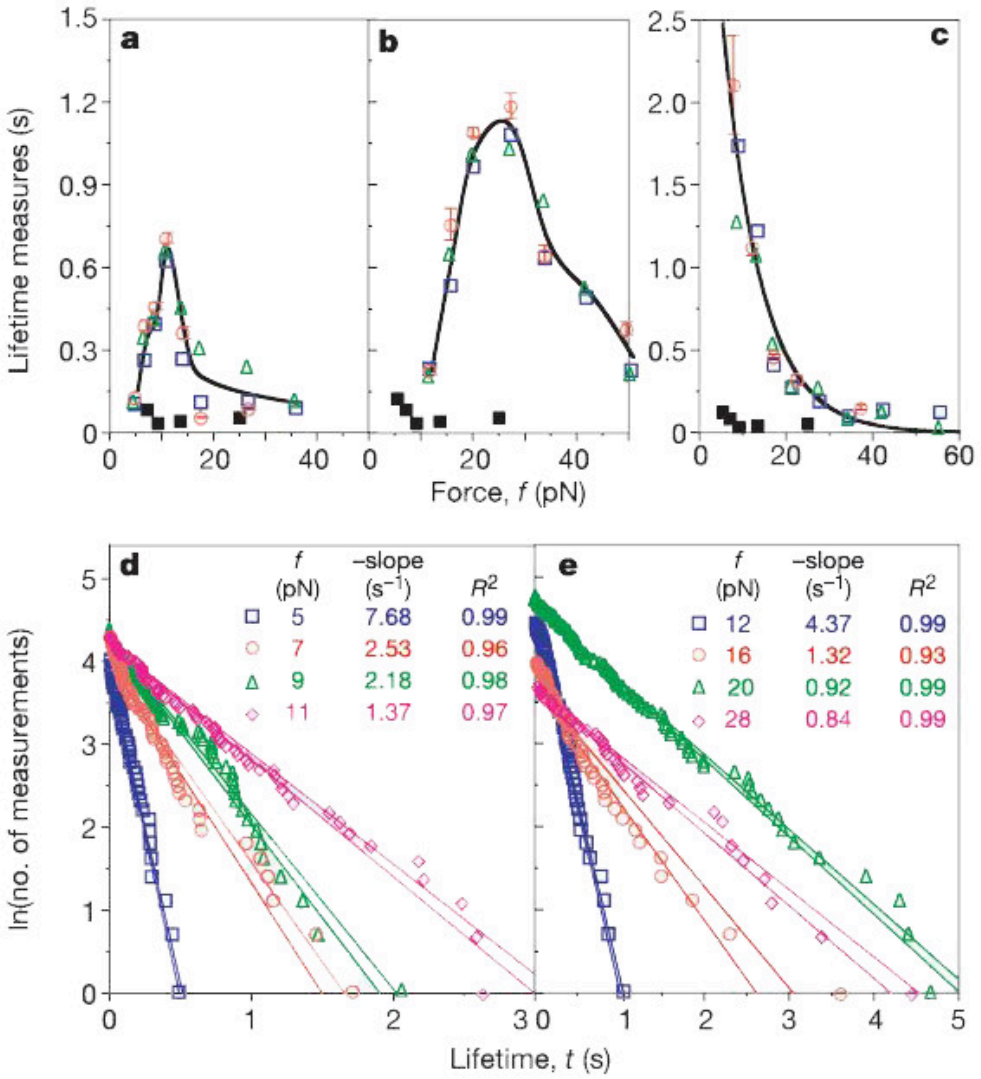


Figure 1: *The mean lifetimes (blue squares) and the standard deviation of lifetimes (green triangles) of P-selectin-PSGL bonds. The displayed dependencies are for the bond between P-selectin with sPSGL-1 (a, d), PSGL-1 (b, e) or G1 (c). Figure by Marshall et al. [1]*

Physical scenarios of catch bonds The biological observations discussed in the previous section show that catch bonds exist. Many theories have been put forward to explain these observations. Figure 2 shows some examples of possible physical scenarios [12]. These explanations have in common that there are two (or more) dissociation pathways: one slow pathway and one fast pathway. Upon increase of the applied force, the fraction of bonds dissociating via the slow pathway increases. Not only the applied force contributes to the dissociation of molecular bonds, but also fluctuations can lead to dissociation. So even if no force is applied, bonds can dissociate.

The molecular bonds showing a catch bonds behavior are complex. This complexity makes it difficult to determine which parts of the molecules are responsible for the catch bond effect. Connecting observations with the theoretical physical scenarios is even more difficult. Creating a simple catch bond, where each part of the molecule has a definite contribution to

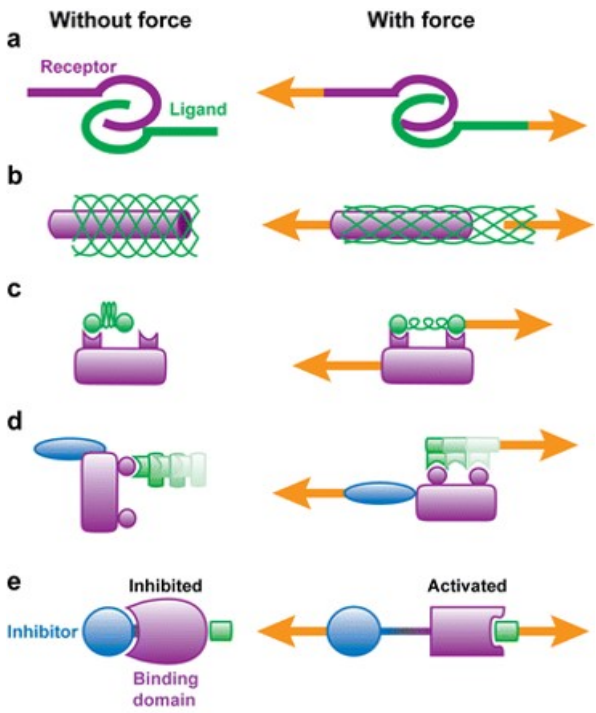


Figure 2: *Examples of conceptual models of catch bonds. In the hook model, ligand and receptor must be brought towards each-other to unbind (a). The deformation model suggests that the ligand deforms under tension, creating a tighter grip (b, c). The sliding-rebinding model assumes that a force aligns regions, resulting in extra connections between ligand and receptor (d). The allosteric model suggests an inhibitor is pulled away by force, increasing the strength of the bond between ligand and receptor (e). Figure by W.E. Thomas et al. [12]*

the catch bond effect, makes it possible to connect experiments and theory. In this way, a humanmade catch bond construct contributes to the understanding of catch bonds.

Materials and other applications Besides contributing to the understanding of catch bonds in nature, humanmade catch bonds also have the potential to be used in materials. For the manufacturing of new technologies, companies want to use optimal materials. At this moment, companies can choose from a wide variety of available materials. Often, improving one aspect of the used material comes at a price of worsening another aspect of the material. In a more optimal situation, each aspect of a material can be tuned to precisely fit the requirements set by the company.

Advancements in material science allow for more and more tunability of properties of matter with earlier unavailable preciseness. The characteristics of these materials can be tuned by changing the involved molecules. For example, the pore sizes in structured carbon materials can be adjusted with hydrothermal carbonisation, emulsion templating, ice templating and organic framework templating [14]. This fine-tuning of pore sizes can, for instance, be used to optimise Li-ion batteries. More recent developments make it possible to fabricate materials with chosen responses to external influences. These programmable materials change their properties following predefined instructions. An example of programmable matter are polymers that change shape when the temperature, pH or used solvent changes [15]. A possible application of this specific example is to use these polymers to create stents to keep the

passageway of blood vessels open.

New advancements make it possible to tune the response of matter to external influences in ways that at first seem counter-intuitive. An example is the temperature response of the DNA coated colloids designed by W.B. Rogers and V.N. Manoharan. In a range of temperatures, colloidal structures aggregate instead of melt with when the temperature increases [13]. Figure 3 gives a visual explanation of the response of this colloidal system to temperature.

Humanmade, tunable catch bonds can be used similarly as above examples to program the lifetime-force dependency of materials. Possible reported advantages of the use of catch bonds in materials include enhanced ductility, toughness, and durability, lower hysteresis and faster relaxation of residual strain than networks of slip bonds [16]. LIterature suggests some specific applications of catch bonds. Examples include using catch bonds as force sensors and using catch bonds for selective drug delivery [17, 18]. So creating catch bonds does not only help explain biological observations but is also potentially relevant for the development of materials and technology.

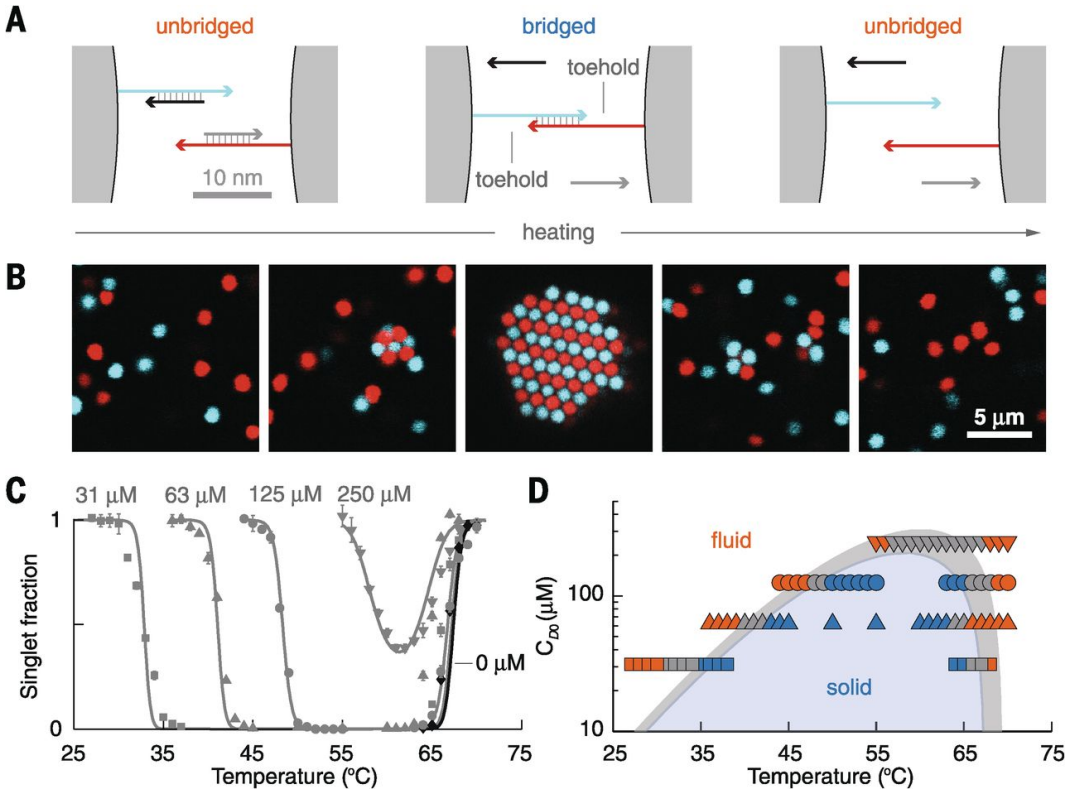


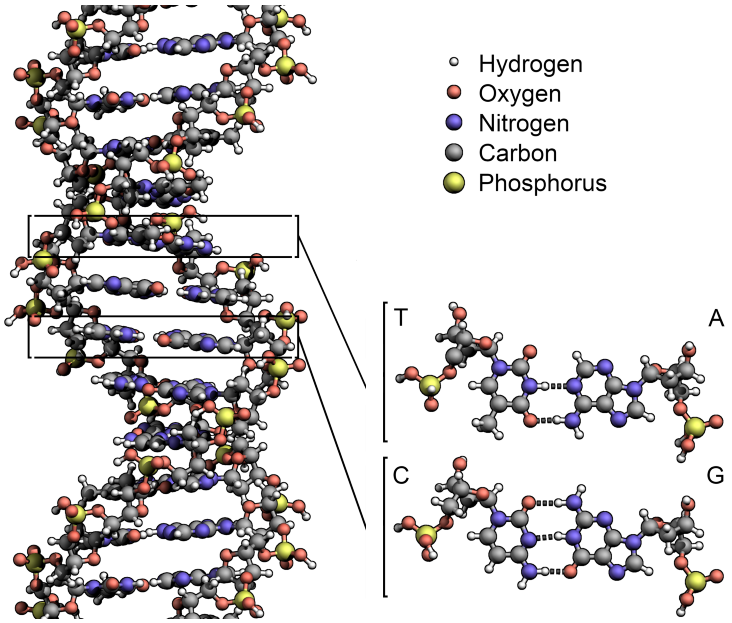
Figure 3: *Inversed singlet-fraction observed for DNA-grafted colloids. Free displacement strands compete with toehold strand for the formation of bonds. For a range of temperatures, the number of connections between toehold strands and free strands decreases (d) while the number of connections between toehold strands increases (a). The result is a decreasing singlet fraction with increasing temperature (c). Figure (b) shows representative confocal fluorescence micrographs. Figure by W.B. Rogers and V.N. Manoharan. [13].*

1.2 DNA as a building block

In the previous sections, we saw that a simple, humanmade catch bond can be used to explain biological observations and to apply catch bonds in technology and materials. The next question is how to create a catch bond. This section explains why DNA is an appropriate choice for a building block.

Deoxyribonucleic acid consists of building blocks called nucleotides. Nucleotides consist of a sugar called deoxyribose, a phosphate group and one of the DNA specific nucleobases. Individual nucleotides form polymers consisting of nucleotides by covalently bonding of deoxyribose of one nucleotide and the phosphate group of the next nucleotide. Each of these DNA strands can bind in its whole to another DNA strand, forming a double helix or double-stranded DNA. Hydrogen bonds between the nucleobases, called base pairs, are the bond connecting the two strands in a DNA duplex.

There are five variations of nucleobases: cytosine (C), guanine (G), adenine (A), thymine (T) and uracil (U). Uracil is usually only found in RNA, so the four common variations of DNA building blocks are C, G, A and T. Due to the location of hydrogen, oxygen and nitrogen in the nucleotides, A binds mostly to T forming two hydrogen bonds and G binds mostly to C forming three hydrogen bonds. So a G/C bond is stronger than an A/T bond. Two strands form a double helix when their bases are (mostly) complementary to each other.



(a) *Schematic picture of the atoms in a DNA double helix.*

(b) *A zoom in on the structure of the four variations of nucleotides, thymine (T), adenine (A), cytosine (C) and guanine (G) and the hydrogen bonds between them.*

Figure 4: *Schematic figure of a piece of double-stranded DNA. Figure adapted from original by R. Wheeler.*

The orientation of base pairs is anti-parallel. So at the end of a duplex, one of the nucleotides has a hydroxyl group (3'-side) and the other one its phosphate group (5'-side). Figure 4 shows a schematic of a DNA helix.

In reality, DNA can have multiple configurations. Long DNA strands have single-stranded and double-stranded regions due to fluctuations. Also, pieces of a strand can bind with pieces of the same strand. DNA can be twisted in both a right-handed and left-handed direction. Furthermore, the distance between bases, the direction of the helix and the stability of DNA changes when the temperature, salt concentration and solvent are changed. As a final example, stretching DNA with high forces can change flexibility and length of DNA.

The designable, specific interaction between bases make DNA a tunable molecule. Many research exists on the force response of DNA. The next sections discuss some of the results of these studies. This high amount makes the of research makes the behavior of DNA under force predictable. The needed force to separate the two DNA strands in small DNA constructs is in the range of tens of pico-Newtons. This force range is available in many experimental methods. These aspects make DNA a suitable molecule to create catch bonds.

Stability and the Nearest-Neighbor model The created DNA catch bonds should be stable. An estimation of the stability of molecules can be made using their free energy. In general, the most stable state is the state with the lowest free energy. The Nearest-Neighbor model is a model used to calculate free energies of single-stranded and double-stranded DNA [22–24]. To calculate the free energy of a DNA strand, this model adds the free energies of pairs of base pairs. Adding contributions of pairs of base pairs is an improvement in comparison to just adding the free energy contributions of individual base pairs since bases influence the neighboring base pairs. These contributions are known as base stacking interactions. The table below shows the free energies used in the Nearest Neighbor model.

The change in free energy per pair of base pairs ($\Delta G^{37^\circ\mathrm{C}}_{\mathrm{stack}}$) is caused by the gain of enthalpy (ΔH) by forming the bond and the decrease of entropy (ΔS). The formation of a bond decreases the entropy because the translational freedom of the involved strands decreases. From the moment the first base pair is formed, two DNA strands cannot move separately anymore. This reduction in translation freedom is the reason why extra initiation energy is added when at least one base pair is formed ($\Delta G^{37^\circ\mathrm{C}}_{\mathrm{initiation}}$). Also, A/T-base pairs often switch between bound and unbound when they are located at one of the ends of a duplex. This also changes the average free energy ($\Delta G^{37^\circ\mathrm{C}}_{AT\,\mathrm{terminal}}$). A last contribution to the free energy is caused by the change in entropy when a DNA duplex is symmetrical. The entropy of a system scales with the number of distinguishable states a system can occupy. The number of configurations of a symmetric DNA duplex is lower than the number of configurations of an asymmetrical duplex, resulting in the final contribution ($\Delta G^{37^\circ\mathrm{C}}_{\mathrm{symmetry}}$) [25].
Combining these contributions gives the total free energy of the DNA duplex:

$$\Delta G^{37^\circ\mathrm{C}}_{\mathrm{total}} = \Delta G^{37^\circ\mathrm{C}}_{\mathrm{initiation}} + \Delta G^{37^\circ\mathrm{C}}_{\mathrm{symmetry}} + \sum \Delta G^{37^\circ\mathrm{C}}_{\mathrm{stack}} + \Delta G^{37^\circ\mathrm{C}}_{AT\,\mathrm{terminal}}. \tag{1}$$

This equation gives the free energy at 37°C. At other temperatures, the free energy is given by:

$$\Delta G^{T} = \Delta H - T\Delta S. \tag{2}$$

In complex DNA structures, the free energy also depends on the arrangement of single-strand and double-stranded DNA. Figure 5 shows some examples. Furthermore, other factors such as

TABLE 1 Nearest-neighbor thermodynamic parameters for DNA Watson-Crick pairs in 1 M NaCl[a]

Propagation sequence	ΔH° (kcal mol^{-1})	ΔS° (e.u.)	ΔG°_{37} (kcal mol^{-1})
AA/TT	−7.6	−21.3	−1.00
AT/TA	−7.2	−20.4	−0.88
TA/AT	−7.2	−21.3	−0.58
CA/GT	−8.5	−22.7	−1.45
GT/CA	−8.4	−22.4	−1.44
CT/GA	−7.8	−21.0	−1.28
GA/CT	−8.2	−22.2	−1.30
CG/GC	−10.6	−27.2	−2.17
GC/CG	−9.8	−24.4	−2.24
GG/CC	−8.0	−19.9	−1.84
Initiation	+0.2	−5.7	+1.96
Terminal AT penalty	+2.2	+6.9	+0.05
Symmetry correction	0.0	−1.4	+0.43

[a]The slash indicates the sequences are given in antiparallel orientation. (e.g., AC/TG means 5′-AC-3′ is Watson-Crick base paired with 3′-TG-5′). The symmetry correction applies to only self-complementary duplexes. The terminal AT penalty is applied for each end of a duplex that has a terminal AT (a duplex with both end closed by AT pairs would have a penalty of +0.1 kcal/mol for ΔG°_{37}).

concentrations of DNA, concentrations of charged and uncharged molecules and the quality of DNA also influence the free energy. SantaLucia et al. made many of these influences quantitative by measuring the differences in free energy. Their papers give more information about these corrections to the free energy ([22–24]).

1.3 Stretching and breaking DNA

Stretching and chain models Polymers, including our designed DNA catch bonds, stretch under the influence of a force. Chain models describe the end-to-end length of polymers under force. Two often used chain models for DNA are the freely jointed chain model (FJCM) and the worm-like chain model (WLCM). The freely jointed chain model assumes rigid pieces of polymer, called Kuhn monomers, are attached with a direction independent of the neighboring monomers. In the freely rotating chain model, the bond angle is constant, but the torsion angle varies randomly like in the freely jointed chain model. In the limit of infinitely small monomer lengths and angles, the freely rotating chain model becomes the worm-like chain model.

Various authors derived chain models specifically for DNA. In this thesis, I use the model by Manghi et al. ([26]). The model used for single-stranded DNA is an extensible, discrete Worm-like chain model. This model is a combination between a freely jointed chain model and a worm-like chain model, with corrections making the model agree better with experiments.

The force-extension curve of double-stranded DNA is slightly more complex due to the

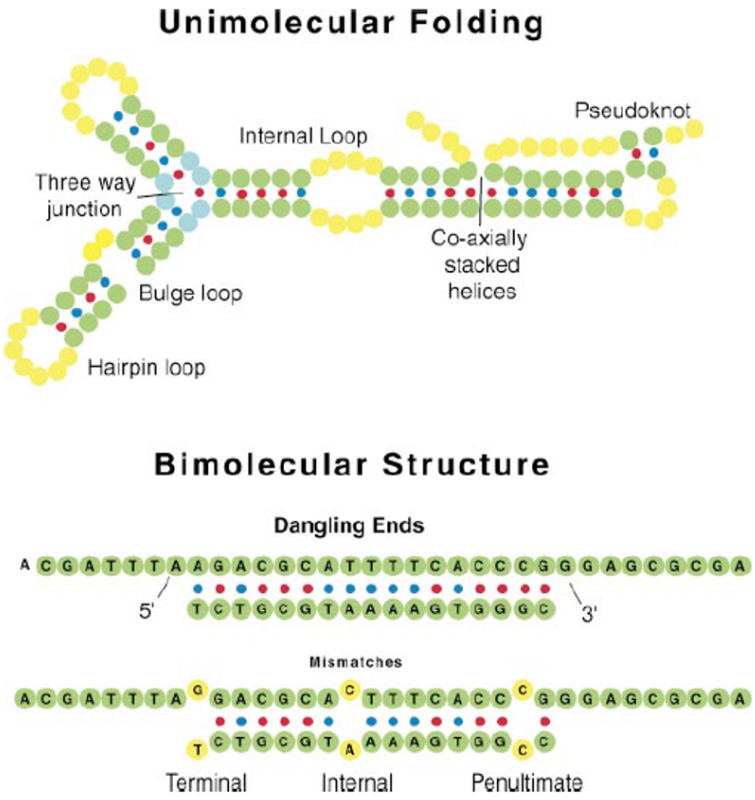


Figure 5: *Visualization of mismatches, loops and other structures influencing the free energy. Figure by SantaLucia and Hicks [24].*

overstretching transition. At forces of about 65 pN, the backbone length of DNA increases significantly. The model used for double-stranded DNA is an Ising-Heisenberg model. This model uses both the force-extension relation of B-DNA and the S-DNA (stretched DNA). The fraction of the strand in each of the states depends on the force. More information about this model can be found in the paper by Manghi et al. ([26]).

Force-induced dissociation of DNA For our DNA catch bonds, we are not only interested in how a force can change the shape of DNA, but also how the connections between (pieces of) DNA strands breaks. Fluctuations randomly open and close base pairs. When all base pairs are open at the same time, a force can move two strands apart so they cannot rebind. So even if the applied force is low, eventually every short DNA complex under force will dissociate. Forces change the likelihood of opening and closing of base pairs. The higher the tensile force applied to a base pair, the lower the average lifetime. The rate equation by T. Strunz et al. describes the dependency of a bond's lifetime when a force is applied [28]:

$$\langle t^f \rangle = 10^{\alpha N - 3} \exp\left(-\frac{f\Delta x}{k_B T}\right) \text{ [s]}; \tag{3}$$

In this equation, N is the number of base pairs in the bound state and $W = \int f \mathrm{d}x = f\Delta x$ is the work done by a constant force. $\langle t^0 \rangle$ is a typical lifetime of the bond when no force is applied. This lifetime depends exponentially on the difference in free energy (in $k_B T$) between

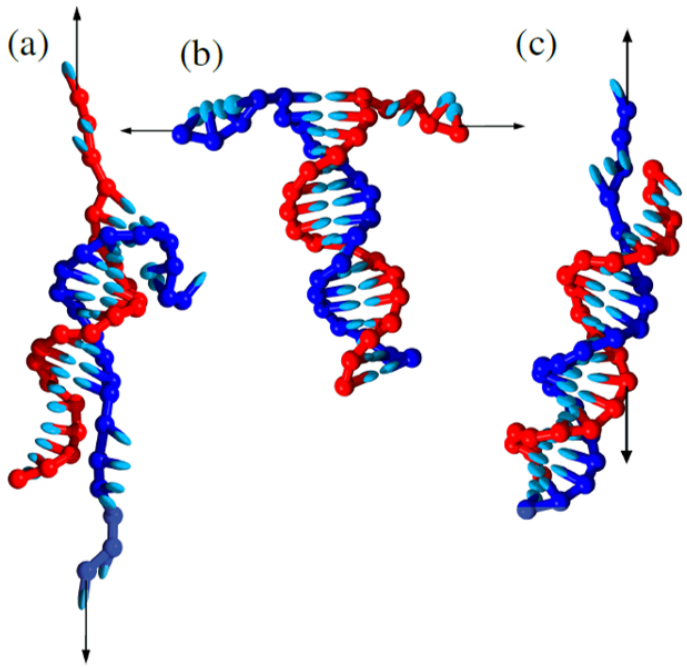


Figure 6: *Schematic representation of shearing DNA (a), unzipping DNA (b) or a combination of both (c). Figure by Mosayebi et al. [27].*

the hybridized and fully dissociated state. Like T. Strunz et al., I approximate this lifetime by $10^{\alpha N-3}$. So only the number of base pairs is taken into account and not the specific bases. T. Strunz et al. found α by fitting this equation with experimental data, giving $\alpha = 0.5 \pm 0.1$. As described earlier, a more accurate prediction of the free energy can be made using the Nearest-Neighbor model [23, 24].

One can open a DNA helix in two ways. Pulling on nucleotides on the same side of a helix is called unzipping DNA. Pulling on opposite ends of a helix is called shear unzipping. 6 shows the difference between these unzipping methods. In both cases, the length of the DNA strand increases with every breaking base pair. In this report, I use the length differences according to T. Strunz et al. [28]:

$$\Delta x_{\parallel} = A + B\,N \tag{4}$$

for shear unzipping and

$$\Delta x_{\perp} = 2\,(A + B\,N)\,. \tag{5}$$

for unzipping.

Shear unzipping DNA increases the end-to-end distance with a constant amount for each unzipped base pair due to the difference in backbone length between double-stranded and single-stranded DNA (B). After the final base pair, a reorientation of the bases results in an additional increase in end-to-end length (A). For unzipping, the end-to-end length increases with the same amount, but now in two directions. In this report, I use $A = 0.7$ nm ± 0.3 nm and $B = 0.07$ nm ± 0.03 nm, as found by Strunz et al.

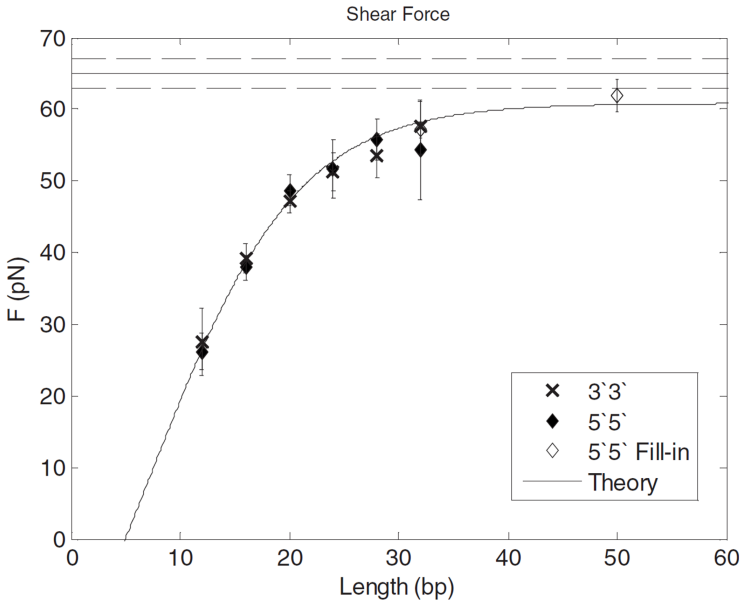


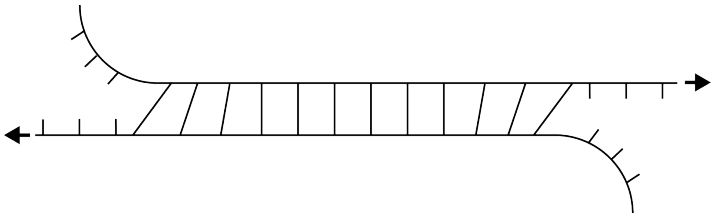
Figure 7: *Graph by K. Hatch et al. [29]. Breaking force as a function of the number of base pairs, for constant force experiments of 1 sec. The solid line represents a fit of the equation by P. de Gennes [30]. The horizontal lines represent the transition to overstretched DNA (from bottom to top 10%,50% and 90%). This data suggests no difference between pulling at the* $3'$*-sides (*$3'3'$*) and the* $5'$*-sides (*$5'5'$*).*

When a helix is unzipped using a shear force, the force is distributed over the base pairs. This stress would be distributed over all base pairs if the backbone of DNA was infinity stiff. However, P. de Gennes and K. Hatch et al. showed that this is not the case [29, 30]. They applied a shear force to DNA duplexes with a varying number of base pairs. De Gennes theoretically derived the shear force needed to unzip a DNA helix as:

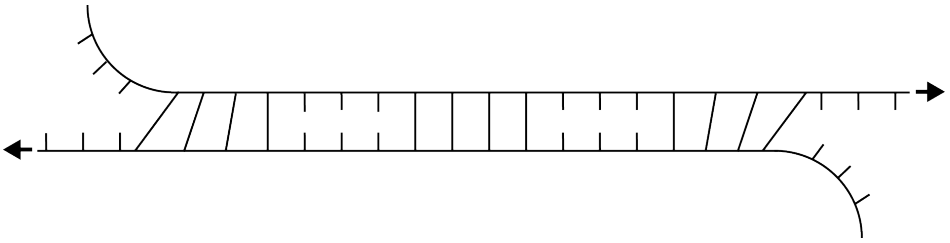
$$\langle F \rangle = f_1 N_N^{eff} = f_1 2 \left(1 + \chi^{-1} \tanh \frac{\chi}{2} N\right). \tag{6}$$

with f_1 the force needed to unzip one base pair and N_N^{eff} the effective number of base pairs over which the force is distributed. χ^{-1} is the adjustment length or the number of base pairs on both sides of the loaded duplex where the force is unevenly distributed over the two strands. K. Hatch et al. showed that for their data the best fit of this equation was found with $\chi^{-1} = 6.8$ base pairs and $f_1 = 3.9$ pN. Figure 7 illustrates the dependence of the dissociation force on the number of base pairs.

The same thing happens when there are multiple regions. The base pairs at the sides experience the highest force. Because there is no force exchange between the strands inside internal loops and hairpins, the distribution of forces in a strand with loops and hairpins is approximately equal to the distribution of forces of attached regions. However, when minimal forces are applied, the same number of base pairs dissociates quicker when regions are not attached. Fluctuations open base pairs next to internal loops faster than base pairs surrounded by other base pairs. This means the average dissociation time of separate regions is lower than the average unbinding time of a helix without internal loops of the same size.



(a) *Shear unzipping a helix without internal loops or hairpins.*



(b) *Shear unzipping a helix with internal loops.*

Figure 8: *Shear unzipping DNA. At each base pair, part of the applied force is transfered to the opposite base and part of the force is transfered to the next nucleotide in the same strand.*

1.4 Research goals and content

Testing the possibility to create a simplistic catch bond is useful to explain biological observations and to make catch bonds potentially applicable in technology and materials. Due to the tunability of DNA and the available knowledge about DNA, DNA is a proper choice for a first molecule. In the used DNA construct, each region contributes to the catch bond effect. This clear functionality makes the behavior of DNA catch bonds interpretable and predictable. In this thesis, we take a closer look at the individual regions in this designed catch bond, as well as the behavior of the catch bond as a whole. The influence of regions and behavior is tested using theory, experiments and simulations. The sections below discuss the results of these tests. The mean question in each of these sections is: "Do the results of this experiment-simulation suggest it is possible to understand and engineer catch bond functionality using DNA?". The individual answers to these questions contribute to answering my main research question "Is it possible to use DNA to create a tunable DNA catch bond?".

This analysis starts with a theoretical analysis of DNA catch bonds. I make a prediction of the behavior of the DNA catch bonds under force using rate equations and free energy calculations. This analysis is used to show that in theory, the DNA constructs behave as catch bonds. After the discussion of the used experimental and simulation methods, chapter 4 shows the results of testing the theoretical model using experiments and simulations. These measurements test specific properties of the DNA catch bonds. The stability of the catch bonds is checked using gel electrophoresis combined with restriction enzymes. These measurements indicate whether the expected ground state of DNA is also the ground state observed in reality.

For the catch bond functionality of the DNA catch bond, it is necessary that a different dissociation pathway is followed for high forces than for low forces. To check whether this is happening in reality, I use optical tweezer experiments. In these experiments, a force is applied to a single DNA construct. Aspects of the resulting force-extension curve show whether the constructs have different dissociation pathways for low and high forces. Next, we

take a look at the increase in lifetime of the catch bonds using simulations. The simulations are performed using coarse-grained molecular dynamics simulation program oxDNA [31]. To speed-up simulations, I use a new sampling scheme developed during my master project, called Stairway Sampling. I attempted to measure the full lifetime-force dependency of DNA catch bonds during this project, but due to problems with the experiments and simulations, I only found a theoretical prediction of this dependency. In each section, I discuss whether the results suggest it is possible to create tunable catch bonds using DNA or not. Finally, I give some recommendations for future research.

2 Theory: A model of DNA catch bonds

In the introduction we learned that a simple, humanmade, and tunable catch bond has the potential to be used in technological applications and to explain biological observations. Also, the previous chapter discussed the advantages of using DNA for this purpose. This section illustrates the proposed way to use DNA to create catch bonds. Also, I make a theoretical prediction whether this DNA catch bond will have an inversed lifetime-force dependency and and how the catch bond could be tuned.

2.1 The DNA catch-bond mechanism

Like the mechanisms discussed in the introduction, DNA catch bonds have two dissociation pathways. At small forces, a low force pathway is more likely while at a high enough force a high force pathway is followed. Figure 9 shows a schematic of the proposed catch bond system. The letters A, B, and H in this figure represent pieces of DNA strand. These pieces of DNA strand are designed using specific sequences of the four nucleotides A, T, G and C. These pieces are designed in such a way that they only bind to specific other pieces of DNA. A only binds with A', B only binds with B', and H only binds with H'. Also, the strands are designed in such a way that they cannot bind with themselves. M.E. Leunissen proposed this particular design.

In equilibrium, the time spent in each of the states is inversely related to the free energy of that state. Each base pair contributes to the free energy. In figure 9, the number at the bottom of each state represents the number of base pairs. For the chosen design, the number of base pairs is maximal in the folded state. So this is the most common state when no force is applied. This does not mean that the DNA strands are and stay in this state forever. Fluctuations open and close base pairs. When all bonds in the handles (H) open due to fluctuations, the strands are no longer connected. This switch of states is the most common dissociation pathway when a weak force is applied.

The backbone of double-stranded DNA has a negative charge [33]. This negative charge results in parts of the DNA strands repelling each other. In the folded state, four directions of double-stranded DNA come together in the middle of the structure. The negative charge of the backbone of each of these pieces pushing them in different directions. Due to the limited flexibility and length of the A/A'-duplexes, it is unlikely that B and B' bind in the folded state. However, also the A/A' parts can be separated by fluctuations, switching the state of the DNA strands to state F_1. In this state, A/A'-duplexes do not keep B and B' apart, so now an extra region can be hybridized (B). The higher number of hydrogen bonds between the two strands stabilizes the connection between the two strands, but also these hydrogen bonds can break. This route is the second possible dissociation pathway. The free energy of the extended state is higher than the free energy of the folded state. So it is unlikely for a bound DNA strand to switch from the folded state to the extended state when no force is applied.

When a high force is applied, the situation changes. Again the catch bonds start in the state with the lowest free energy, so the folded state. However, forces increase the likelihood for hydrogen bonds to break. The effect of an applied force is larger on the base pairs in the hairpins than on the handles. Every dissociating base pair in the hairpins results in an extension about twice as big as the extension gained by shear unzipping a base pair in the handles. So energetically, unzipping the hairpins is more favorable than unzipping the

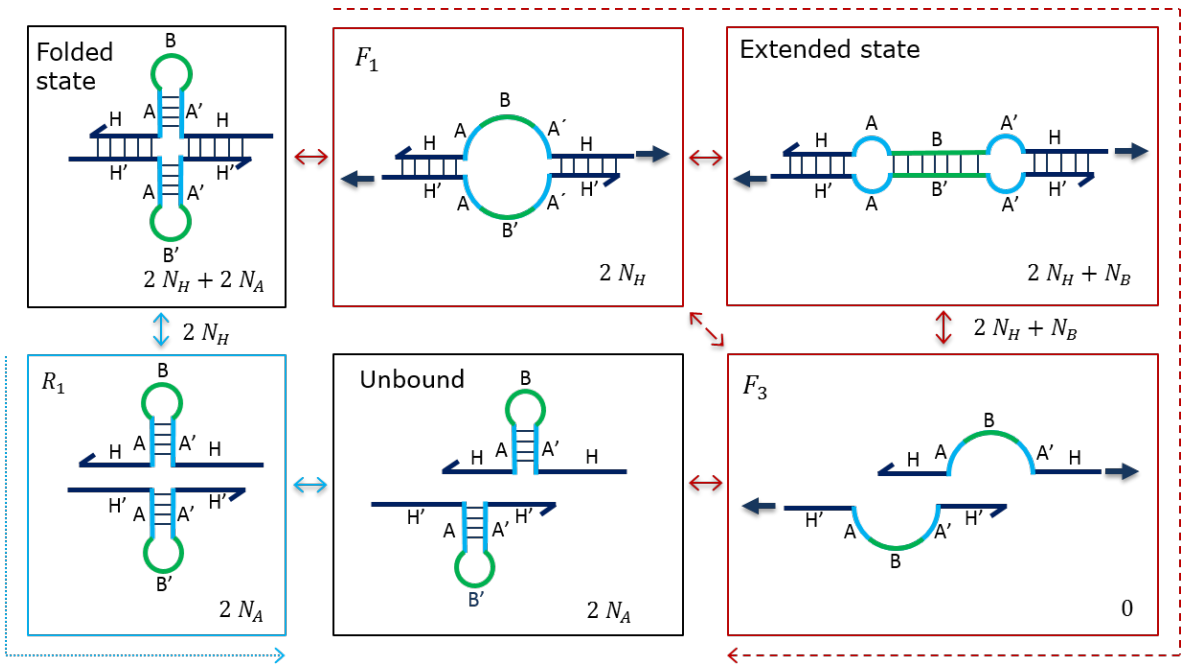


Figure 9: *Schematic representation of the states and pathways of DNA catch bonds. The number at the bottom of each state represents the number of bonds in that state. Nucleotides in A can form base pairs with nucleotides in A', B with B' and H with H'. When a small force is applied, a bond can break by following the blue dotted R pathway. This pathway requires* $2N_H$ *base pairs to lose the connection between the DNA strands. When a high enough force is applied to break the A/A'-bonds faster than the H/H'-bonds, the red dashed F pathway is followed. This pathway requires* $2N_H + N_B$ *bonds to lose the connection between the DNA strands.*

handles. So, although the lifetime of the H/H'-bonds decreases, the lifetime of the A/A'-bonds decreases quicker, increasing the probability of switching from the folded state to the extended state. After this transition, the stabilizing region aligns, allowing for quick hybridization. To break the two DNA strands apart, now the bonds between the two handles and the extra region must break.

2.2 DNA Catch Bond requirements

The idea described above can be made more quantitative by looking at the different pieces of the DNA catch bond. In this section, we take a closer look at these pieces and derive requirements for their lengths. Also, this section explains how changing the number of bases in each of these pieces tunes the DNA catch bond.

The first requirement is the stability of the construct. The folded state must be the most common state when no force is applied, and individual regions should stay bound within experimental timescales. Secondly, there must be a critical force, where the most common pathway switches from the low force pathway to the high force pathway. As a part of this high force pathway, the B/B'-region should bind. Finally, the dissociation time of the high force pathway should be higher than the dissociation time in the low force pathway. Here, we look quantitatively at these requirements using DNA models introduced in the introduction.

The requirements for the catch bond effect are translated to requirements for the number of base pairs in each of the pieces of DNA using the results of this analysis.

Free energies and lifetimes are used to explain the requirements and tunability of DNA catch bonds. The free energy contribution to the unbinding time is calculated using the model by T. Strunz et al., which assumes equal contributions to the free energy of each base pair [28]. This is an approximation to reality, since many factors influencing the free energy. The reasoning in this section could be repeated more accurately for specific constructs using more accurate free energy calculations by using the Nearest-Neighbor model [23, 24].

Also, we assume that the dissociation time of DNA decreases exponentially with force, in the same way as in the model by T. Strunz et al. [28]. The model only differentiates between shear unzipping and unzipping DNA. So possible effects caused by the specific shape of the unzipped DNA strands and the attached pieces of DNA strands do not change this dependency. Also, rebinding of regions is not taken into consideration. Both of these assumptions can be removed by calculating the rates of every step in the unbinding process for each of the possible unbinding pathways. The used model assumes that the end-to-end length increase upon unzipping depends on the length of the DNA sequences, but not on the applied force. In reality, DNA under a higher force extends more, also resulting in a slightly higher difference in length upon unzipping.

When an extra base pair is added to a long strand of DNA, the effect on the unbinding force during force-ramp experiments is smaller than for short DNA strands. De Gennes described this effect with an effective number of base pairs [30]. This effect is included in the calculations below by using the effective number of base pairs in the equations.

Stability A catch bond effect is only possible if the designed structured is stable when no force is applied. In this thesis, we look at single catch bond constructs. When no force is applied, the average unbinding time equation by T. Strunz et al. [28] reduces to:

$$\langle t^0 \rangle = 10^{\frac{1}{2}N_N^{eff}-3} \text{ [s]} \tag{7}$$

So pieces of strand with 8 base pairs dissociate on average in about 10 seconds. Each region should have a measurable effect, so the timescales involved should be larger than the timescales involved in other processes during experiments. The time needed to reach a constant force in optical tweezer experiments is a couple of seconds, so 8 base pairs are minimally required. So translating this discussion to the first requirement for the number of base pairs:

$$\begin{aligned} 2N_H &\geq 8; \\ N_A &\geq 8. \end{aligned} \tag{8}$$

For structures with multiple catch bonds in parallel, lower number of base pairs can be used. Besides the stability, also the initial state of the DNA construct is important. The effect of state switching is maximal when the folded state is the most common state before a force is applied. The folded state is the most common state when there are more (or stronger) base pairs in the folded state than in the extended state, so:

$$2N_H + 2N_A \geq 2N_H + N_B, \tag{9}$$

or

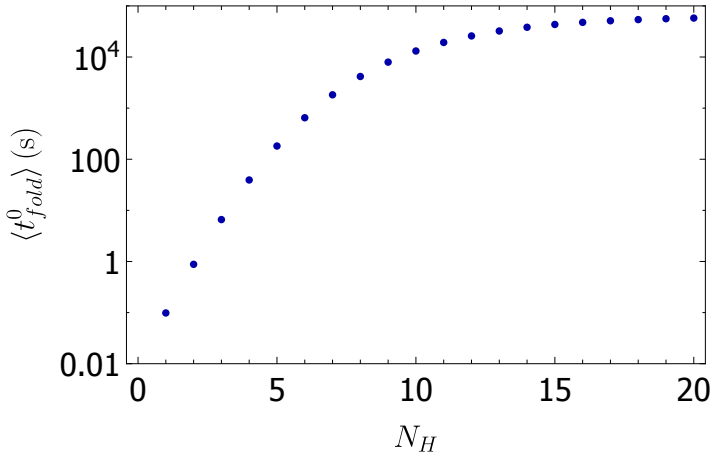


Figure 10: *The average lifetime of DNA catch bonds can be tuned by varying the length of the handles.*

$$2N_A \geq N_B. \tag{10}$$

When the free energy of the hairpins and the stabilizing region are comparable, a catch bond effect is still possible. In that case, part of the constructs is in the folded state and part of the constructs in the extended state. The constructs starting in the folded state would be responsible for an increase in the average dissociation time under force.

So the ground state of the DNA catch bonds is the folded state. In this state, the two DNA strands are connected via the two handle regions with lengths N_H. The average lifetime of the DNA catch bonds when no force is applied can be tuned by changing the length of the handles:

$$\langle t^0_{fold} \rangle = 10^{\frac{1}{2}N^{eff}_{2N_H} - 3} \text{ [s]}\,. \tag{11}$$

Figure 10 shows a plot of the lifetimes at zero force.

Force-induced state-switching Now our theory defines when the initial folded state is stable, we look at the conditions needed to allow switching to the extended state under the influence of force. When no force (or a low force) is applied to a construct, the handles should break before the hairpins open. Without force, the hairpins and the handles both open by fluctuations. The unbinding time of these regions can be approximated using the number of base pairs in these regions:

$$\begin{aligned} \langle t^0_H \rangle &= 10^{\frac{1}{2}N^{eff}_{2N_H} - 3} \text{ [s]}\,; \\ \langle t^0_A \rangle &= 10^{\frac{1}{2}N^{eff}_{2N_A} - 3} \text{ [s]}\,. \end{aligned} \tag{12}$$

So the requirement that $\langle t^0_H \rangle < \langle t^0_A \rangle$ can be translated to

$$N_A > N_H. \tag{13}$$

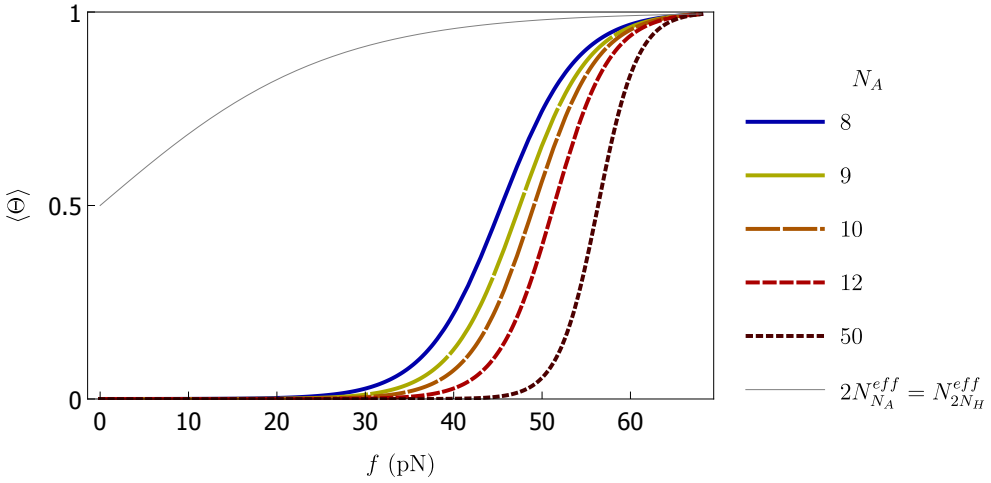
Above a critical force, the hairpins should open before the handles break. Under force the lifetime of regions of DNA is given by:

$$\langle t^f \rangle = 10^{\frac{1}{2}N_N^{eff}-3} \exp\left(-\frac{f^{eff}\Delta x}{k_B T}\right). \tag{14}$$

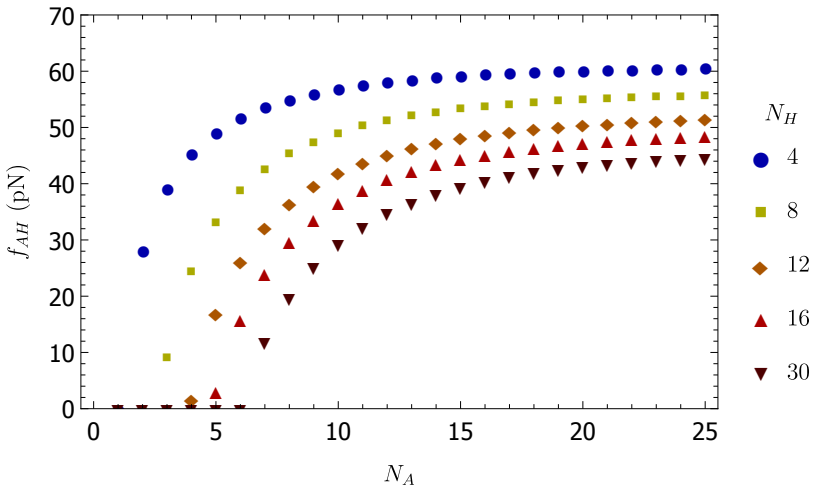
with

$$\Delta x = \left(0.7 + 0.07\, N_N^{eff}\right) \text{ [nm]} \tag{15}$$

for shear unzipping and



(a) *Fraction of constructs switching to the extended state as a function of force for* $N_H = 8$.



(b) *Critical force for a variation od=f* N_H *and* N_A.

Figure 11: *The higher the force applied to a DNA catch bond construct, the higher the chance that the hairpins open before the handles break. The force where* $\Theta = 0.5$ *increases with* N_A. *The higher* N_H, *the lower the critical force for big values of* N_A.

$$\Delta x = 2\left(0.7 + 0.07\, N_N^{eff}\right)\ [\text{nm}] \tag{16}$$

for unzipping DNA, as discussed in section 1.3. The increase in end-to-end distance per unzipping base pair is higher for unzipping than for shear unzipping. This difference makes it more energetically favorable to unzip a strand than to shear unzip a DNA strand. The double hairpin in a Holiday-junction has a higher energy barrier, but the same Δx as a single hairpin.

A shear force applied to a construct consisting of multiple regions distributes over the regions. In the folded state, both of the handles are loaded equally and feel an effective force of $f_{2N_H}^{eff} = \frac{1}{2}f$. When one of the handles breaks, the full force targets the other handle. So the dissociation time of this second handle is much smaller than the dissociation time of the first handle. Thus the total dissociation time is approximately the dissociation time of the first handle. The full force loads the hairpins, so $f_{2N_A}^{eff} = f$.

At the critical force, the average time needed to switch to the extended state, by opening of the hairpins, is equal to the time needed break the catch bond, by opening the handles. So to find the critical force, we just need to set both times equal and solve for the force:

$$\begin{aligned} \langle t_A^{f_{AH}^c} \rangle &= \langle t_H^{f_{AH}^c} \rangle \\ 10^{\frac{1}{2}2N_{N_A}^{eff}-3} \exp\left(-\frac{f_{AH}^c\, \Delta x_A}{k_B T}\right) &= 10^{\frac{1}{2}N_{N_H}^{eff}-3} \exp\left(-\frac{\frac{1}{2}f_{AH}^c\, \Delta x_H}{k_B T}\right); \end{aligned} \tag{17}$$

$$f_{AH}^c = 68 \cdot 2\frac{2N_{N_A}^{eff} - N_{N_H}^{eff}}{4N_{N_A}^{eff} - N_{N_H}^{eff}}\ [\text{pN}]\,. \tag{18}$$

In this equation, the Δx are chosen up to the maximum energy of the barrier, so $\Delta x_A = 2 \cdot N_{N_A}^{eff}\, 0.07$ nm and $\Delta x_H = N_{N_H}^{eff}\, 0.07$ nm. This equation is only correct for $2N_{N_A}^{eff} > N_{N_H}^{eff}$, because only than $\langle t_{N_A,0}^0 \rangle > \langle t_{N_H}^0 \rangle$s.

The maximal critical force according to this equation is 68 pN for $N_{N_A}^{eff} \gg N_{N_H}^{eff}$. This maximal force can only be reached for very short handles. For high N_A, the force is distributed over part of the bases, limiting the increase in critical force. For $N_H = 10$, the maximal critical force is for instance 54 pN instead of 68 pN.

At the critical force, half of the constructs will switch from the folded to the extended state before they break. Around the critical force, the number of constructs switching to the extended state increases with force. The fraction of constructs that switch to the extended state depends on the likelihood that the hairpins open before the handles break. This fraction (Θ) can be estimated using the average rate of occurrence of these events:

$$\begin{aligned} \Theta &= \frac{\nu_A}{\nu_A + \nu_H}; \\ &= \frac{10^{\frac{1}{2}\left(N_{N_H}^{eff} - 2N_{N_A}^{eff}\right)} \exp\left(\frac{f\left(\Delta x_A - \frac{1}{2}\Delta x_H\right)}{k_B T}\right)}{10^{\frac{1}{2}\left(N_{N_H}^{eff} - 2N_{N_A}^{eff}\right)} \exp\left(\frac{f\left(\Delta x_A - \frac{1}{2}\Delta x_H\right)}{k_B T}\right) + 1} \end{aligned} \tag{19}$$

Figure 11 shows the fraction of constructs in the folded and the extended state for varying hairpin lengths. This figure clarifies that for every handle length, the critical force can be tuned by varying the length of the A-regions.

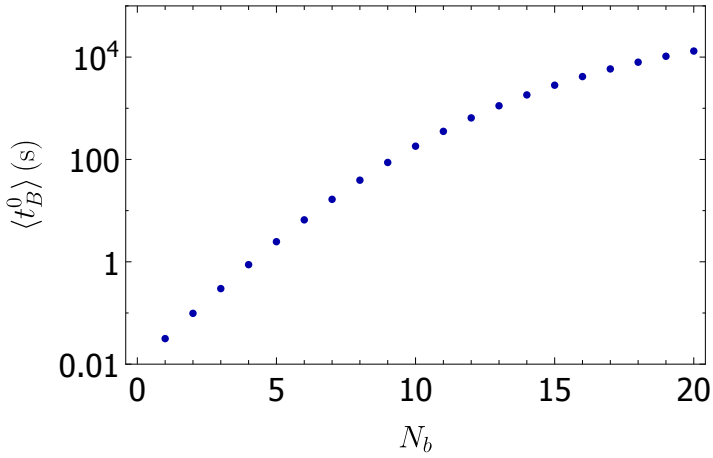


Figure 12: *The relaxation time from the extended state to the folded state, after reducing the applied force to* 0 *pN.*

Hybridization stabilizing duplex In the previous section we saw that in theory, force can induce a switch from the folded state to the extended state. After this switch, the B-region should hybridize. After the hairpins unzip the large internal loops stretch and the B and B' regions align, decreasing the time required for hybridization. An internal loop separates the B-region and the handles, so there should be enough bases in the stabilizing region to hybridize by itself and have a significant lifetime.

After the rupture of a catch bond, it is desired that if a catch bond rebinds, it rebinds in the folded state. When a catch bond dissociates from the folded state, the hairpins stay connected, and rebinding will be in the folded state. If a catch bond dissociates from the extended state, the hairpins of each of the strands will quickly rebind due to their complementarity and proximity.

A catch bond in the extended state returns to the folded state when you stop applying a force. The time needed for this relaxation depends on the length of the B-region according to

$$\langle t_B^0 \rangle = 10^{\frac{1}{2}N_{N_B}^{eff}-3} \text{ [s]}\,. \tag{20}$$

This time is plotted in figure 12.

To limit the experimental time, it would be convenient if this time is not too long. In that case, multiple subsequent experiments can be performed without the need to wait for long times until the catch bond relaxed to the folded state. At the same time, the B-region should be long enough to bind and affect the dissociation time. Using a minimal lifetime in the range of tens of seconds and a maximal lifetime in the range of hundreds of seconds results in a minimum of 8 bases and a maximum of 12 bases. When enough relaxation time is available, the length of the B-region can be longer to increase the catch bond effect. So to summarize:

$$8 \leq N_B. \tag{21}$$

Increasing lifetime After the B-region hybridizes, the lifetime of the DNA catch bond should be longer than before the hybridization. So the dissociation time of the extended state under force should be higher than the dissociation time of the catch bond in the folded state

under a lower force. To find the effect of the extra connecting region between the two DNA strands, again the average dissociation times can be compared:

$$\begin{aligned}\langle t^f_{fold}\rangle &= 10^{\frac{1}{2}N^{eff}_{N_H}-3}\exp\left(-\frac{f^{eff}_{2N_H}\,\Delta x_H}{k_BT}\right);\\ \langle t^f_{ext}\rangle &= 10^{\frac{1}{2}N^{eff}_{N_{H/B}}-3}\exp\left(-\frac{f^{eff}_{H/B}\,\Delta x_{H/B}}{k_BT}\right).\end{aligned} \tag{22}$$

In the folded state, the force is again distributed over the two handles, resulting in an effective force at each of the handles of $f^{eff}_{2N_H} = \frac{1}{2}f$. In the extended state, the effective force depends on the loaded region. The two handles are loaded equally. They are closer to the source of the force, so receive part of the force according to their relative size. The extra, middle region is loaded with the rest of the force:

$$\begin{aligned}f^{eff}_H &= \frac{1}{2}\frac{N^{eff}_{2N_H}}{N^{eff}_{2N_H+N_B}}f;\\ f^{eff}_B &= \left(1-\frac{1}{2}\frac{N^{eff}_{2N_H}}{N^{eff}_{2N_H+N_B}}\right)f.\end{aligned} \tag{23}$$

When one of the regions opens, the force redistributes over the other regions. These regions dissociate much faster due to the higher force. When the B-region is much smaller than the H-regions, the B-region opens in a very short time. The rest of the time is spent opening the handles. In that case, the lifetime of the DNA catch bond in the folded and the extended state would be approximately equal. So for a catch bond effect, the extra double-stranded region should be at least comparable in size to the size of the handles.

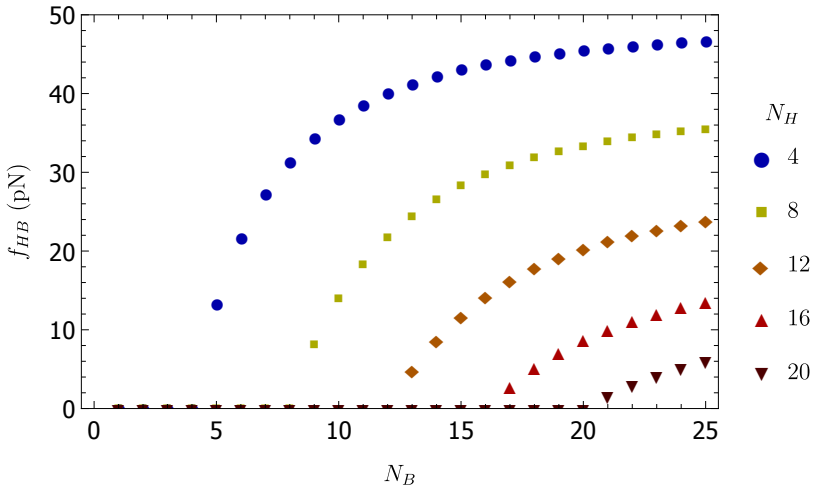
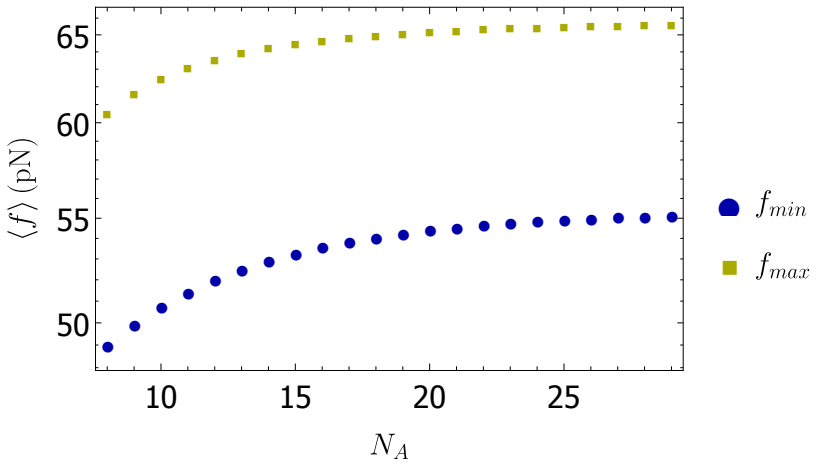


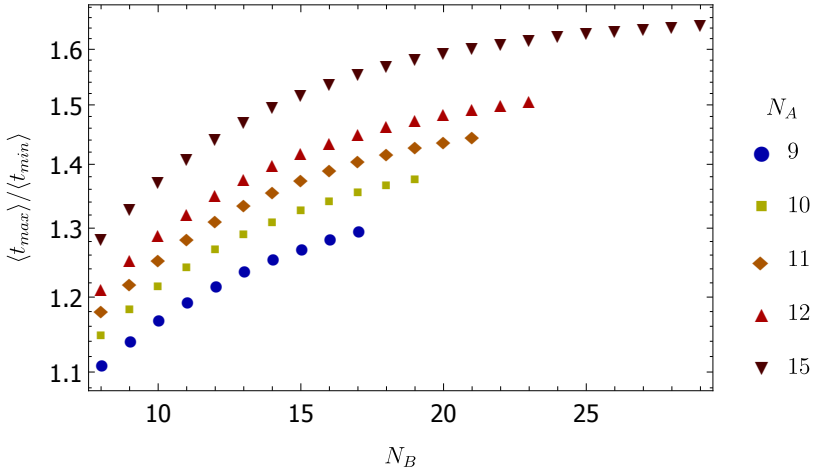
Figure 13: *Critical force F_{HB}, varying N_H and N_B. Above this force, the added lifetime as a result of the slower unzipping of the handles becomes larger than the added lifetime as a result of the unzipping of the B-region.*

If $N_B \approx N_H$, the total time can be approximated with the time needed to open one of the regions. In this case, it is likely that one of the handles breaks first since they are loaded more than the extra region. The increase in lifetime is in this case caused by the decrease in effective force applied to the handles.

If the $N_B > N_H$, the handles will dissociate first since they are loaded with the highest force and less stable. For high forces, the lifetime can again be approximated by the lifetime of the first dissociating handle. For low forces, the lifetime of the fully loaded extra region is higher than the lifetime of the handles under a smaller load. In that case, the total lifetime can be approximated by the lifetime of the extra region. The critical force separating these two regimes is given by:



(a) *The forces corresponding to the minimum and maximum of the lifetime-force curve ($N_B = 12$).*



(b) *Relative difference in dissociation time between the maximum and minimum of the lifetime-force curve.*

Figure 14: *Numerical solutions of the difference in time and force at the maximum and minimum of the lifetime-force dependency. The higher N_B and N_A, the larger the relative difference in time. In these figures, $2N_H = 8$.*

$$
\begin{aligned}
\langle t_{B'}^{f_{HB}^c} \rangle &= \langle t_{H'}^{f_{HB}^c} \rangle; \\
10^{\frac{1}{2}N_{N_B}^{eff}-3} \exp\left(\frac{f^c - HB\,\Delta x_B}{k_B T}\right) &= 10^{\frac{1}{2}N_{N_H}^{eff}-3} \exp\left(\frac{f_H^{eff}\,\Delta x_H}{k_B T}\right); \\
f_{HB}^c &= 68 \cdot \frac{2N_{2N_H+N_B}^{eff}\left(N_{N_B}^{eff} - N_{N_H}^{eff}\right)}{2N_{2N_H+N_B}^{eff} N_{N_B}^{eff} - N_{2N_H}^{eff} N_{N_H}^{eff}}\,[\mathrm{pN}].
\end{aligned}
\tag{24}
$$

Above this critical force, the average time can be approximated with the needed time to open the first handle. Below this force, the average time is approximately equal to the time needed to break the extra region. The maximal possible critical force is 68 pN and is reached when $N_B^{eff} \gg N_{N_H}^{eff}$. Figure 13 shows the critical forces as a function of N_A and N_B.

The fraction of constructs following the high force pathway is described by equation 19. So combining the equations above and equation 19, we find the average dissociation time:

$$
\langle t_{tot} \rangle = \Theta \langle t_{ext} \rangle + (1 - \Theta) \langle t_{fold} \rangle. \tag{25}
$$

The next section shows figures of this average lifetime-force dependency. To make an estimation of the size of the increase in lifetime, we compare the maximum and the minimum of this dependency. Figure 14 shows the numerical solutions of f_{min} and f_{max} and t_{max}/t_{min}.

As expected, the increase in lifetime is larger for bigger B-regions than for smaller B-regions. The relative difference also increases with N_A. This can be understood by looking at the dependency on N_A of f_{min} and f_{max}. Both forces increase with N_A, but f_{min} slightly faster than f_{max}. In both the folded and the extended state the average unbinding time decreases with force. Due to the higher dependency of f_{min} than f_{max} on N_A, t_{min} decreases more with N_A than t_{max}. Also, higher values of N_A allow for higher values of N_H.

An increase in lifetime is only achieved when state switching results in a maximum. Due to sequential unzipping, this is only possible when

$$
N_H \leq 8. \tag{26}
$$

For $N_H = 7$ and $N_N = 8$, N_A and N_B should be sufficiently large. So under this condition, also the final requirement for a catch bond effect is satisfied. We saw before that the lifetime when no force is applied can be tuned by varying N_H and that the critical force can be tuned by varying N_A. In this section, we saw that when N_H and N_A are chosen, the increase in lifetime with force can be tuned by varying N_B.

Summary and interpretation The analysis at the previous pages shows that in theory, all requirements for the DNA catch bonds to work as catch bonds can be satisfied:

- Stability and correct formation;
- Force induced state-switching
- Hybridization of the extra, stabilizing duplex;
- Increase of lifetime after the state-switching and formation of the extra duplex;

The requirements for a single DNA catch bond can be summarized as follows:

$$\begin{aligned} 4 \leq & N_H \leq 8; \\ 8 \leq & N_A; \\ 8 \leq & N_B \leq 2N_A. \end{aligned} \tag{27}$$

For $N_H = 7$ and $N_N = 8$, N_A and N_B should be sufficiently large to give an increase in lifetime. Catch bonds combined in parallel can have even smaller numbers of base pairs due to collective stability.

We also saw that the behavior of the catch bonds changes with the length of individual regions. The lifetime when no force is applied can be tuned by varying the length of the handles (N_H). The critical force, around which the lifetime increases, can be tuned by varying the length of the hairpins (N_A). The increase in lifetime can be tuned by varying the length of the additional region in the extended state (N_B). So by setting first N_H, than N_A and than N_B, each of these important characteristics can be set. Also, all of these variables can, in theory, be tuned over a wide range of values, especially when multiple catch bonds are used in series or parallel. This simple tunability makes the catch bond potentially relevant for technology and new materials.

The values presented in this theoretical analysis should be interpreted as rough predictions and not as exact numerical solutions. As discussed, not all contributions to the free energies are taken into account. Also, the analysis does not include different pathways at the scale of single bases and the dissociation time is assumed to depend exponentially on the energy barriers and force, without taken details in the energy landscape. Taking into account these details would change the results quantitatively. The reasoning in this section could be repeated more accurately for specific constructs using more accurate free energy calculations by using the Nearest-Neighbor model [23, 24] and by including all possible pathways in the calculations.

2.3 Lifetime-force dependency

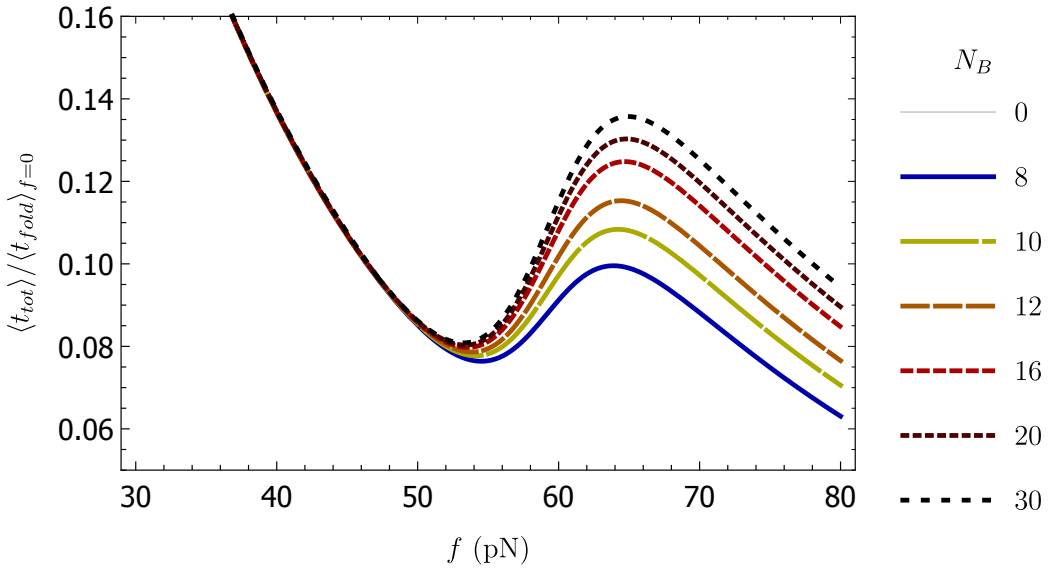
Catch bonds have the property that their lifetime increases with force over a range of forces. For DNA catch bonds, the force-lifetime dependency of is described by

$$\langle t_{tot} \rangle = \Theta \langle t_{ext} \rangle + (1 - \Theta) \langle t_{fold} \rangle. \tag{28}$$

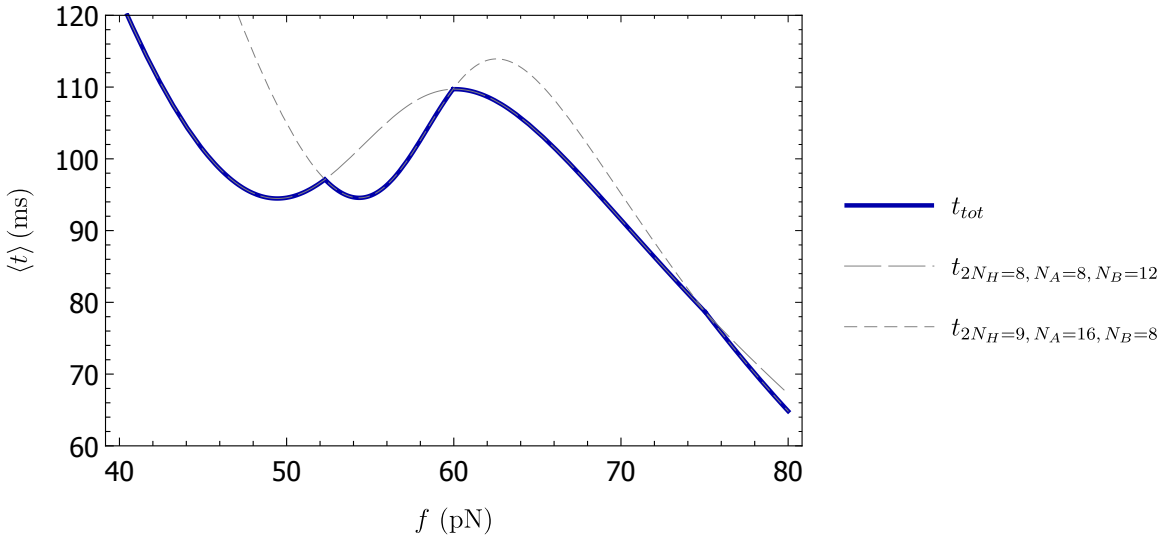
This equation makes it possible to plot the lifetime as a function of force. Figure 15 shows some examples of this dependency. In a range of forces, the lifetime increases with force.

Combining DNA catch bonds (or combining catch bonds and slip bonds) in series or parallel results in an even more tunable lifetime force dependency. When two bonds are used in series, the full connection breaks when one of the connections breaks. So the lifetime is approximately equal to the lifetime of the bond with the smallest lifetime. Figure 15 shows an example of such a combination. When two bonds are combined in parallel, both bonds must break before the full construct break, so the lifetime is at least the lifetime of the strongest bond. By combining bonds, the lifetime-force dependency can be tuned to meet technological or material requirements perfectly.

So in theory, our DNA constructs have a catch bond dependency. The rest of this thesis discusses how this theoretical prediction is tested using experiments and simulations. Based on the arguments in this chapter, we designed several DNA sequences, which are together with the experimental and simulational techniques discussed in the next chapter.



(a) *Lifetime-extension of DNA catch bonds, varying N_B. In this plot, $N_H = 4$ and $N_A = 16$.*



(b) *Two DNA catch bonds in series.*

Figure 15: *Lifetime-force dependency for a catch bond and a combination of two catch bonds. By using variations of catch bonds and slip bonds in parallel and in series, the lifetime-force landscape can be tuned.*

3 Methods: Experiments and Simulation

The previous section shows that in theory, our DNA construct is a catch bond. To check whether the theoretically predicted behavior is also observed in reality, I performed multiple experiments and simulations. This section discusses the used experimental and simulational methods.

3.1 DNA synthesis and colloid functionalisation

DNA synthesis To be able to perform experiments on DNA the DNA sequences must be synthesized. *Eurofins Genetics* and *Integrated DNA Technologies* synthesized the oligos used in our experiments using phosphoramidite chemistry [34].

For the synthesis, the first nucleotide of each strand is attached to a solid support. Then, the oligonucleotides are synthesized in four stages. These four stages are repeated for as many times as the number of bases in a sequence, with the addition of one base to the strands in each cycle.

In the first stage, the already present nucleotide is altered in such a way that it has an available binding side at the 5'-side of the nucleotide. For this, a protecting dimethoxytrityl group is removed (detritylation). In the second step, new nucleotides are added to the solution. These nucleotides couple with the free 5' hydroxyl group (coupling). During the coupling step, not all free hydroxyl groups react with the new bases. To reduce the number of truncated sequences, the DNA sequences that did not react are capped using acetylation (capping). However, a small percentage of strands stay uncapped, resulting in wrongly sequenced oligonucleotides. In the last step, an unstable phosphite triester binds to the 5'-side of the sequence. After oxidation, this results in nucleotide capped by a dimethoxytrityl group. Now the process can be repeated to add another base to the sequence. Figure 16 shows the DNA synthesis cycle.

After the synthesis process, you have a solution with the sequenced DNA. There are, how-

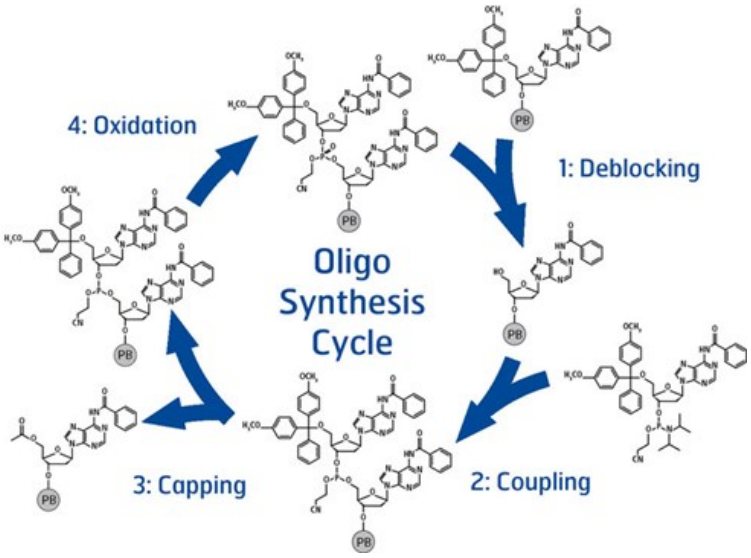


Figure 16: *Oligo synthesis cycle. Figure from the Biologio website.*

ever, still incorrectly sequenced oligos. The DNA is purified using PAGE purification to reduce the amount of incorrectly sequenced strands. In this purification method DNA sequences are filtered based on their size and conformation using polyacrylamide gel electrophoresis. The purity, so the relative amount of DNA strands without undesired bases, is at least 85% after this purification process. The purified DNA is checked using Optical Density measurements and Matrix Assisted Laser Desorption/Ionization Time of Flight measurements. With these measurements, the mass of the strands can be checked to estimate the purity.

I designed the specific sequences using NuPack [35]. Their design tool makes it possible to specify the desired interactions of DNA strands and their optimization protocol suppresses undesired interactions. For the final design choices, I took into account the melting temperature of the wanted and unwanted structure, the number and relative percentage of off-target interactions, the possibility for sliding-rebinding interactions and the expected interaction with restriction enzymes. These characteristics can be checked by calculating the free energies of (pieces of) the DNA complex and by comparison of pieces of sequences.

The first part of each of the designed sequences for the optical tweezer experiments is a piece of strand complementary to the strand that is used to attach the sequence to beads. The following part is a poly-A spacer. In each of the corners of the Holiday-junction have a T that is not part of any of the complementary regions. Possible spatial interactions limit base pairing at the middle of the Holiday-junction. The middle base pairs are chosen to be non-complementary to neighboring bases to eliminate an effect of these middle base pairs on the free energies. In this way, we know these base pairs do not form hydrogen bonds, instead of being uncertain about their most common state.

Appendix A shows all tested sequences in the gel electrophoresis, centrifugal microscopy, optical tweezers and DNA coated colloid melting experiments. I choose these sequences in such a way that they are suitable for the experimental setup and that the researched property of the catch bonds can be measured optimally. It should be noted that this also means that for many of the constructs a catch bond effect is not expected according to our theory described in the previous chapter.

Colloid functionalisation In some of the performed experiments, we functionalized colloids by attaching DNA catch bonds to the surface of the colloid. By applying a force to these colloids, indirectly a force is applied to the connecting catch bonds. I use DNA coated colloids in optical tweezer experiments, centrifugal microscopy experiments and melting temperature experiments. This section discusses how we functionalized these beads.

To create the functionalized colloids used in the optical tweezer experiments, we first create a solution with the full used DNA complexes. The total sequences are not only the DNA catch bonds but also two spacers of 5004 base pairs, attached to both sides of the catch bond. These spacers are used to increase the distance between the DNA catch bonds and the beads, to limit the effect of the beads during experiments. At the end of one of the two spacers, we attach digoxigenin (DIG). A convalent bond connects the other spacer with biotin. These attached molecules are later used to connect the DNA complexes to colloids using a biotin/NeutrAvidin-bond and a DIG/anti-DIG-bond.

The separate DNA sequences are attached to each other using a thermocycler, forming a complex consisting of the two sides of the DNA catch bond and the two spacers. A twenty base pair long sticky end forms the connection between the spacers and the DNA catch bond. The strength of this connection is improved by a DNA ligase enzyme catalyzing the formation of

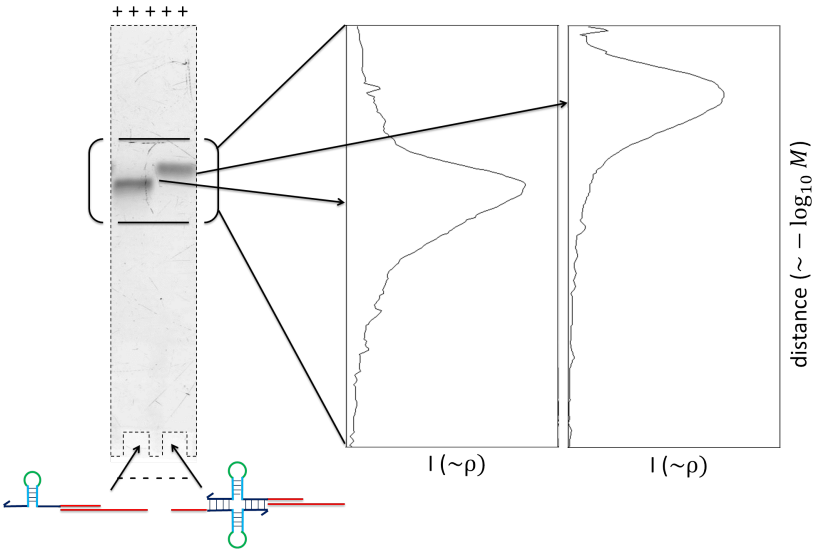


Figure 17: *Visual explanation of the used gel electrophoresis method. A solution of both strands (X and Y) or one strand (X) are added in the same concentrations to a gel. An electric potential is applied over the gel forcing the negatively charged DNA through the gel. After a set time, you make a fluorescence image. Computer analyses of this image shows the intensities at specific locations in the gel.*

covalent phosphodiester linkages between the sticky ends and the ends of the double-stranded DNA. The two sides of the DNA catch bond should not be ligated, because it should be possible to dissociate catch bonds during the experiments. Such a ligation is blocked using five single-stranded poly-A bases.

Now we have the DNA solution, the DNA constructs can be attached to the colloids. For this purpose, anti-digoxigenin (anti-DIG) antibodies were attached covalently to carboxyl-functionalized polystyrene microspheres using the crosslinker carbodiimide [36]. A DIG/anti-DIG bond form the connection between one side of a catch-bond and a colloid and a biotin/NeutrAvidin bond a second connection. The resulting pair of colloids, connected via DNA tethers is used in the optical tweezer experiments. The preparation of the functionalized beads for the centrifugal microscopy experiments is similar. For the centrifugal microscopy experiments, I used magnetic beads instead of polystyrene beads. Also, instead of using spacers at the ends of the DNA catch bonds, a DNA NanoSwitch is attached to both sides of the DNA catch bond. Nanoswitches are discussed later in this chapter. Instead of a ligated connection between the different parts of the DNA construct, a long DNA duplex is used (60 base pairs). The connection between the magnetic bead and the Nanoswitch is a Biotin/Steptavidin bond.

The strength of the bond between the colloids and the DNA was tested using multiple experiments with a DNA construct consisting of only two attached spacers. During these experiments, it was possible to increase the force multiple times to the over-stretching force without breaking the construct. These experiments suggest the connections are strong enough to perform experiments up to this force.

3.2 Gel electrophoresis: Analyzing stability and structure

Gel electrophoresis makes it possible to confirm that the DNA catch bonds form in the expected shape. In this section, I discuss the gel electrophoresis method. Gel electrophoresis is a technique to estimate the size and shape of charged molecules. When you use this technique, you first cast a gel, in our case using either Acrylamide (6%-12%) or Agarose (0.7%-2%). Next, you add a solution with the analyzed molecule to one side of the gel. Then you apply a voltage over the gel for a limited time, in our case varying between 30 minutes and 90 minutes.

The idea of gel electrophoresis is that a molecule moves through a gel due to the electric potential applied over the gel. In our case, using DNA, the molecules move from the negative side of the gel to the positive side due to the negative backbone charge of DNA. Friction with the gel limits the velocity of the molecules. This friction depends on the shape and size of the molecule. So after running a gel for some time, small particles are located further away from their starting point than bigger molecules. Comparing these distances with distances travelled by molecules with known sizes an estimation can be made of the size of the measured molecules. During our gel electrophoresis experiments, we used DNA ladders with duplexes of $10 + 5n$ and $50 + 50n$ base pairs.

The final step is finding where the molecules are at the end of the gel electrophoresis experiment. For DNA, a fluorescent dye can be attached to DNA strands before or after the gel electrophoresis experiment. DNA can be synthesized in such a way that a fluorescent dye attaches at a specific location. The advantage of this method is that the intensity of a measured fluorescence signal scales directly with the number of strands. Also, you know that the signal corresponds to the (part of) the molecule with the fluorescent dye. The other option is staining the gel. In this case, fluorescent particles attach to the backbone of DNA after the gel electrophoresis experiment. An advantage of this method is that the DNA molecules do not have to be altered before the experiment. Since more dyes attach to larger DNA strands, it is not possible to directly compare signal intensities to find relative amounts of strands with this method. Figure 17 explains the used procedure.

Restriction enzymes Combining gel electrophoresis with restriction enzymes makes it possible to check the structure of the DNA catch bonds in more detail. The used enzymes cut the sequences of DNA at specific locations, only when those specific sequences are double-stranded. Figure 18 shows the place where the enzymes cut the DNA. The used enzymes are AseI and MluCI to cut the hairpin and BsRBI to cut the extended construct in the B/B'-domain (enzymes from *New England Biolabs*). The A/A'-region is only bound in the folded state. So if AseI or MluCI cuts the structure, this confirms the folded state forms (or another structure where A/A' is bound). In the same way, BsRBI shows whether the extended state forms.

In our experiments, we added the tested DNA strands to a CutSmart buffer (*New England Biolabs*). This buffer is optimized for the functionality of the used restriction enzymes. We leave the sample with CutSmart Buffer, DNA and the used restriction enzyme for one hour. During this hour, the enzymes can interact with the DNA. Directly after this hour, the DNA is analyzed using gel electrophoresis.

It should be noted that it is possible that a DNA strand switches from for the folded state to the extended state and back during the hour of activity of the restriction enzymes. So if all of the DNA strands interacted, not necessary all strands are in the state where they interact

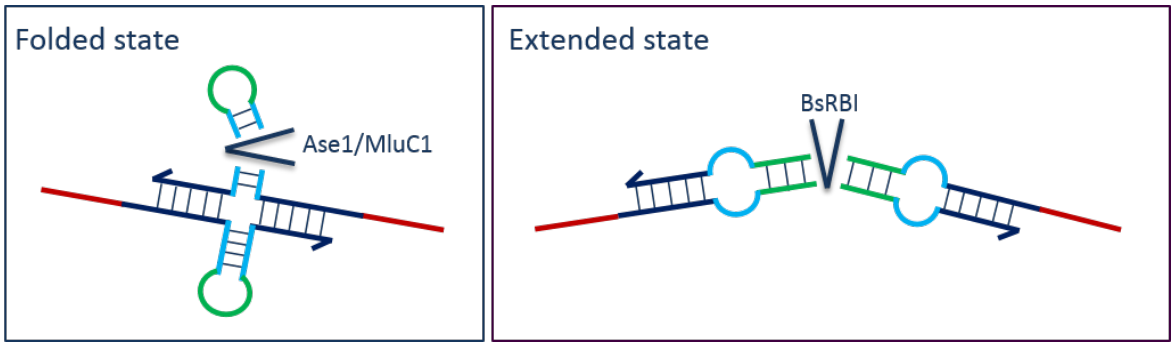


Figure 18: *Schematic of where the constructs get cut by the enzymes AseI, MluCI and BsRBI. For the used construct with $N_H = 10$, the size of the complete construct is 234 bases. After the cut by AseI/MluCI, the sizes of the pieces of DNA are 20 and 214 bases. After the cut by BsRBI, the two pieces of DNA contain 117 bases.*

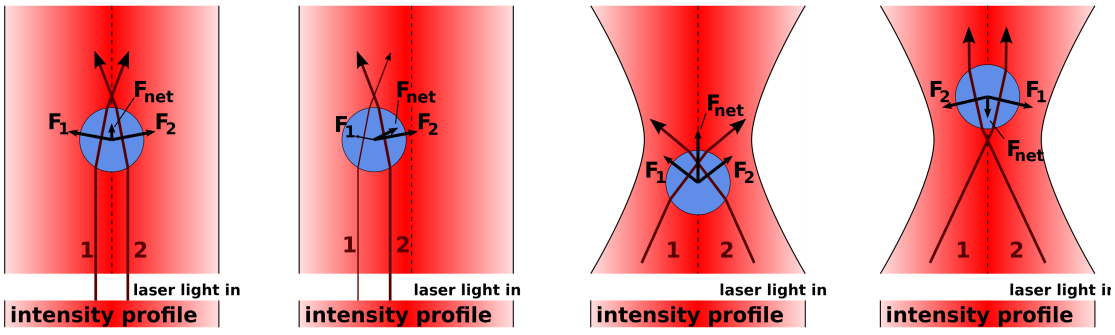
with the restriction enzymes all the time.

3.3 Optical Tweezers: Stretching and dehybridization of DNA

After testing the stability of DNA catch bonds, I tested the functionality of parts of the bonds. In these tests, DNA catch bonds are stretched and broken using optical tweezers. This section describes the used method.

When a ray of light hits a surface, part of the light passes through the surface, and part of the light scatters from the surface. During both of these processes the momentum of the photons in the ray changes. A change in momentum of the surface object balances this change in momentum. The backscattering of photons results in a scattering force away from the source of the light beam. The refraction of light results in a refraction force, which is the force used in optical tweezer experiments.

A ray of light through a sphere bends in the direction of the center of the sphere (see



(a) *Force on microsphere by ray refraction of an inhomogeneous unfocused laser.*

(b) *Force on microsphere by ray refraction of an inhomogeneous focused laser.*

Figure 19: *Force on microsphere by ray refraction. Intensity differences result in a net force in the direction of the center of the laser beam. If the laser is focused, the refraction of light rays result in a net force in the direction of the focus of the light beam. Figure by R. Koebler.*

figure 19a). In accordance with Newton's third law, this results in a force moving the center of the sphere closer to the light ray. At the same time, the scattering force pushes the bead away from the source of the light ray. If a focused laser is used, a bead also experiences a force towards the focus of the beam (see figure 19b). This force balances the scattering force slightly behind the focus point of the light beam. So a focused light beam can trap microspheres.

The force applied to trapped objects depends on the intensity of the laser and the location of the trapped object relative to the focus of the laser. For small displacements of microspheres from the center of the laser beam, the force varies linearly with displacement [37]. So by measuring the location of a microsphere over time, the force applied to the microsphere can be measured.

For the results in this thesis, I used optical tweezers to trap colloids coated with DNA catch bonds. Figure 20 shows the used setup. In this setup, a blue LED illuminates the

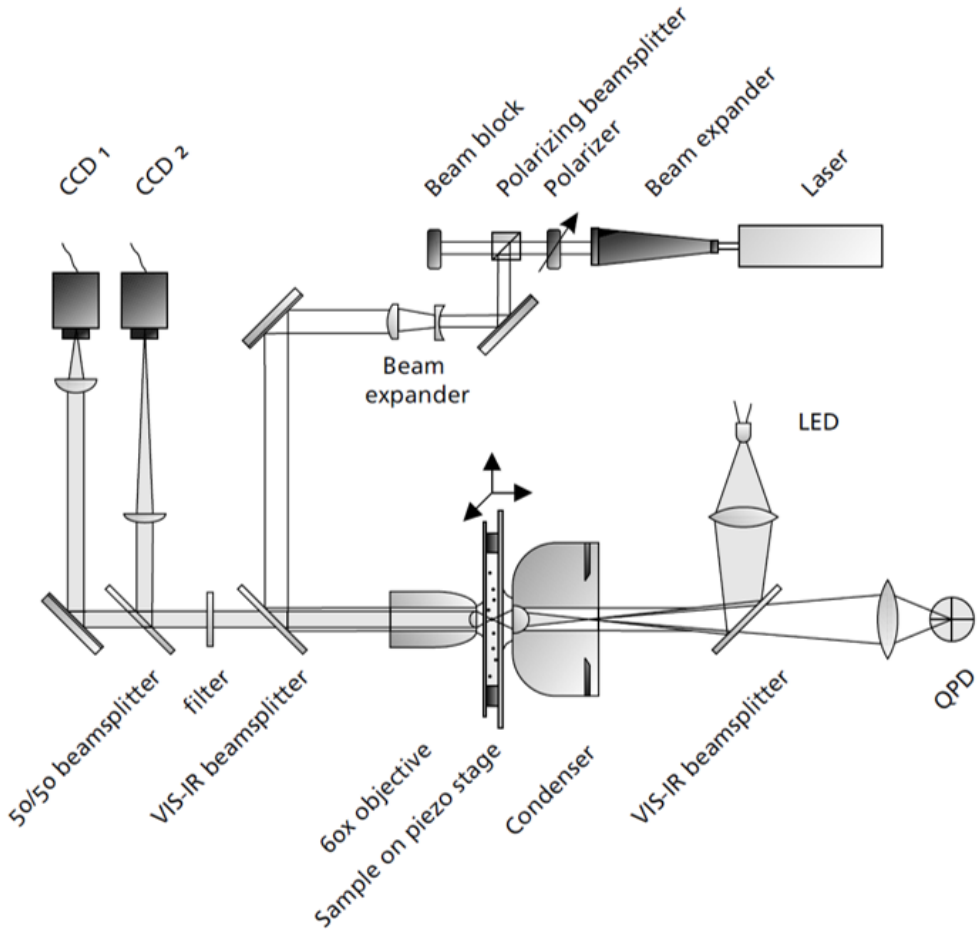


Figure 20: *Schematic drawing of the optical tweezer setup. The sample can be moved relative to the laser using a piezo stage. Two cameras are used to image the sample. A* $Nd:YVO_4$ *laser with a wavelength of 1064 nm is focused in the sample. Two cameras and a quadrant photo-diode track the locations and forces of trapped microspheres. Figure by Dr. R.G.H. van Leeuwen.*

sample, two cameras give live footage of the sample and a $Nd:YVO_4$ laser is focused inside the sample. This laser is used to trap the beads. Together with the cameras, a quadrant photodiode measures the location of the trapped bead relative to the focus point of the laser. A small displacement causes a change in the refraction of the laser light, which results in a higher illumination of one of the quadrants. The difference between the voltage of each of the quadrants can be used to calculate the displacement of the trapped microsphere, which can be used to calculate the force exerted on the microsphere.

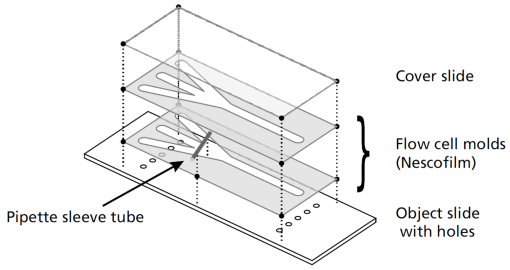
Before each experiment, I prepare three solutions. The first solution contains only PBS (Phosphate-buffered saline). This solution is used to clean the sample slide. The second and third solution contains DNA catch bond coated colloids suspended in PBS and NeutrAvidin coated colloids suspended in PBS. A flow channel as shown in figure 21b is attached to Teflon tubes connected to syringes containing the solutions. First, I add NeutrAvidin coated colloids to the sample. Once the laser traps one of these colloids, I attach the trapped colloid to a micropipette. After cleaning the channel, I add the DNA catch bond coated colloid to the sample. After removing the other colloids from the sample using the PBS solution, the sample is ready for an experiment.

The sample can be moved relative to the laser focus using a piezo stage. In this way, the colloid attached to the pipette can be moved relative to the colloid trapped by the laser. The two colloids are moved towards each other until they connect. Now a force can be applied to a connecting tether by moving the two beads apart. By monitoring the displacement of the trapped colloid, the applied forces can be measured and controlled.

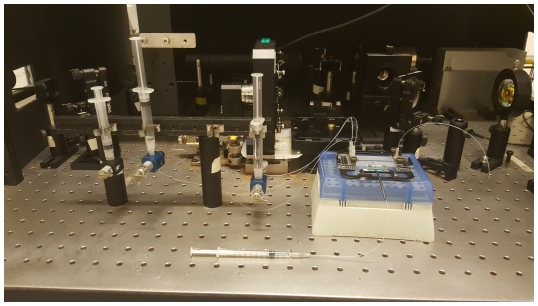
Calibration of the Optical Tweezers setup The force applied at each moment in time is measured using the signal of the quadrant photodiode. The force on the beads can be approximated by Hookean spring [38]:

$$F = k_x \left(x - x_0\right). \tag{29}$$

For small displacements of the microsphere, the displacement scales linearly with the measured quadrant voltage differences in each direction [37]. So:



(a) *Schematic of the component of the flow cell. Figure by Dr. R.G.H. van Leeuwen.*



(b) *Flow cell connected to reservoirs in syringes.*

Figure 21: *Experiment preparation. First, the flow cell is assembled (a). Then the assembled flow cell is connected to the suspension containing syringes using tubing (b). Finally, the flow cell is placed inside the microscope (background b).*

$$V_x = S_x \left(x - x_0\right), \tag{30}$$

with V_x the difference in voltage in the x-direction and S_x the sensitivity of the detector. Combining these equations gives:

$$F = \frac{k_x}{S_x} V_x. \tag{31}$$

Although all beads are similar in shape and content, their response to the laser varies slightly. To make sure the correct force is measured during each experiment, I determined $\frac{k_x}{S_x}$ before each new experiment. We did this using the power spectrum method [38]. This method uses the displacement fluctuations of a trapped microsphere under the influence of the laser. The relation between the power spectral density is approximately Lorentzian [39]:

$$PSD(f) = \frac{k_B T}{\pi\gamma \left(f_c^2 + f^2\right)}. \tag{32}$$

In this equation, $PSD(f)$ is the power spectral density as a function of the frequency of the oscillations (f), k_B is the Boltzmann constant, T the temperature and $f_c = \frac{k_x}{2\pi\gamma}$ is the corner frequency. By fitting this Lorentzian through the power spectral density-frequency, the ratio $\frac{k_x}{S_x}$ can be found. On average, $k_x = 1.63 pN/\mu m$ and $S_x = 1.63 V/\mu m$.

Experimental procedure Now the system is set up and calibrated, the experiment can be started. Using a Labview program, the applied force (or distance between the beads) can be controlled over time, while measuring the distance between the beads. The full setup was initially built by R. van Leeuwen [36]. Options include applying a constant force, increasing or decreasing the force at a constant rate or increasing or decreasing the distance between the beads at a constant rate. During experiments, I save the calibration data and location of

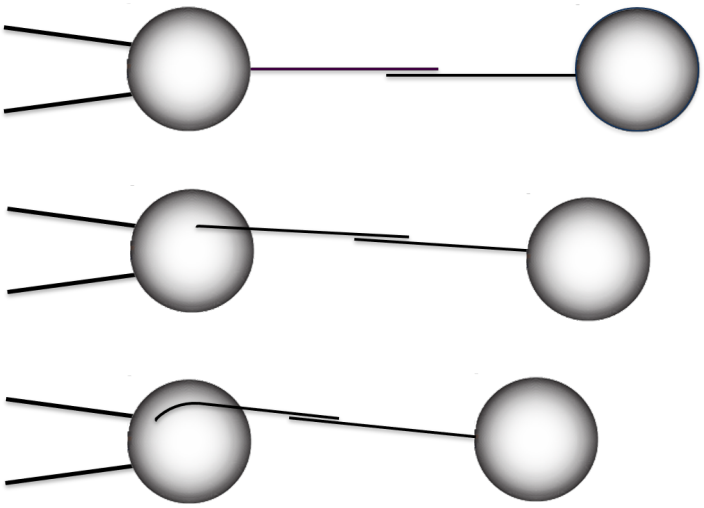


Figure 22: *Different locations of the tether under force. Attachment location for the left (pipette trapped) colloid is equal to the initial binding location. The right colloid rotates in such a way that the tether aligns with the force. Long (5024 kbp) spacers are attached on both sides of the DNA catch bond to suppress interaction with the colloids.*

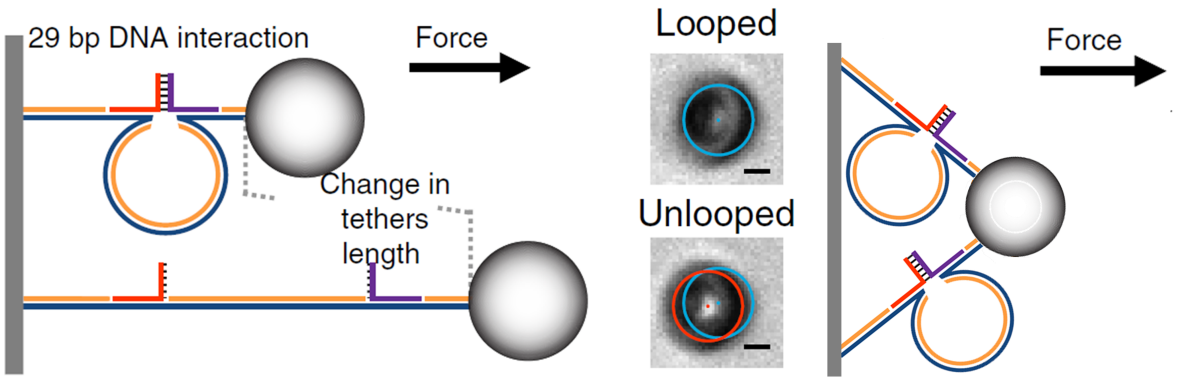
the centers of the beads and the quadrant photodiode voltages over time. This data is later used to plot force-time, extension-time and force-extension curves.

Experimental force-extension curves are shifted on the extension axis to align with the theoretical force-extension curves. The required data shift is a result of the unknown absolute distance between the ends of the DNA strands. This distance is unknown because the size of the colloids varies slightly and because the attachment location of the tethers to the colloids is unknown.

The tethers forming the connection between two beads can attach anywhere at the surfaces of these beads. When a force is applied, the optically trapped bead rotates in such a way that the force is minimal, so with the tether attached at the location closest to the other bead. The other colloid, trapped by the micropipette, can not rotate freely, so the location where the tether attaches initially is also the location during the experiments (see figure 22). So the end-to-end distance of the stretched DNA construct can vary relative to the distance between the centers of the beads. Although the absolute end-to-end distance cannot be used to check what is happening, it is possible to use the shape of the force-extension curve to check that experiments are performed correctly. Also, the overstretching force can be compared with the experimental expected overstretching force to check the experimental calibration.

3.4 Nanoswitches: Increasing experimental throughput

A single lifetime-force dependency requires the results of many dissociation experiments. Nanoswitches can increase the experimental throughput. This section explains what Nanoswitches



(a) *Schematic of closed and open Nanoswitch.*

(b) *Difference in camera view between a closed and opened Nanoswitch.*

(c) *If the distance between the Nanoswitches on the surface is less than the length of a Nanoswitch, multiple Nanoswitches can attach to the same bead.*

Figure 23: *Illustrations of Nanoswitches adapted from Yang et al. [40]. The Nanoswitch can switch between a closed state where the tested DNA strands are hybridized and an open state where the tested DNA strands are single stranded. when a force is applied to the displayed bead and the Nanoswitch opens, the position of the particle shifts in the direction of the force. The scale bars in figure (b) have a length of 1 μm.*

are and how they can be used.

Nanoswitches are long DNA strands that connected surfaces and particles. A tested DNA construct attaches at two locations in this Nanoswitch. If the two sides of a tested construct are attached to the Nanoswitch and each other, the Nanoswitch is closed. If the connection between the two sides of the tested construct open, the Nanoswitch is open. Figure 23 illustrates the idea.

The advantage of the using a Nanoswitches is that rupture experiments can be repeated quicker. The Nanoswitch keeps the two sides of the tested construct near each other, decreasing the average time until the two sides of the DNA construct find each other. So with Nanoswitches, you can perform a higher number of bond lifetimes measurements in the available time.

Nanoswitches also offer a way to check that the measured transitions are the transitions that one wants to measure. Besides the bond between the DNA strands, also the attachment between DNA and particles or a surface can break. When Nanoswitches are used, such a dissociation results in the particle moving more than the length of the Nanoswitch, so more than a desired dissociation. In this way, it is possible to distinguish wanted transitions from unwanted transitions. Also, the direction of the movement of beads can be used to check that the beads are not connected to multiple tethers (see figure 23c).

The Nanoswitches used in our experiments are created by linearizing an M13 virus. This DNA strand has a backbone of around 2.5 μm. Using 120 complementary strands with 60 nucleotides and one with 49 nucleotides block nearly the full M13. However, four pieces of 60 bases are left single-stranded. Two of these pieces are used to attach the strand to a surface and a particle. Each of the other two single-stranded pieces is used to attach the tested DNA strands. When a Nanoswitch opens during an experiment, the bead moves to a new position and slightly out of focus. This shift can be used to measure bond lifetimes.

3.5 Centrifugal Microscopy: Parallelization of experiments

Using Nanoswitches, we can already measure more bond lifetimes within a certain amount of time. Centrifugal microscopy makes it possible to perform many measurements in parallel, making it possible to do even more measurements in this time. This section discusses how

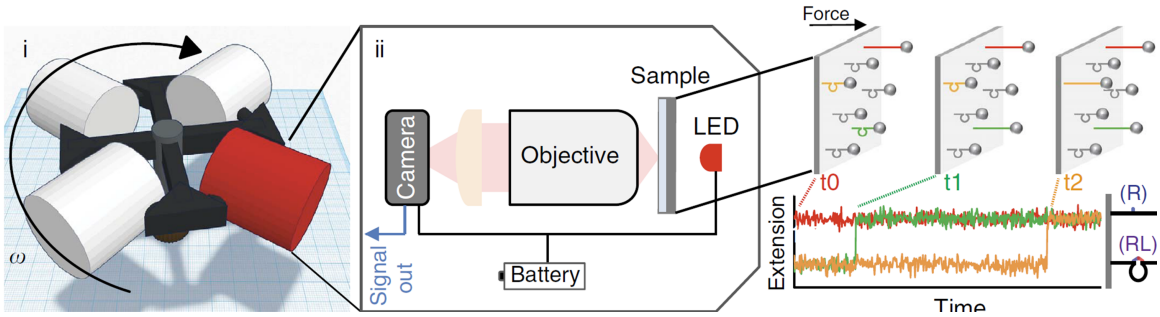


Figure 24: *The centrifugal microscopy setup. Rotating a sample applies a centrifugal force to the particles. The position of these particles is registered by a camera. By registering sudden changes in position, the lifetime of bonds can be measured. Figure by Yang et al. [40].*

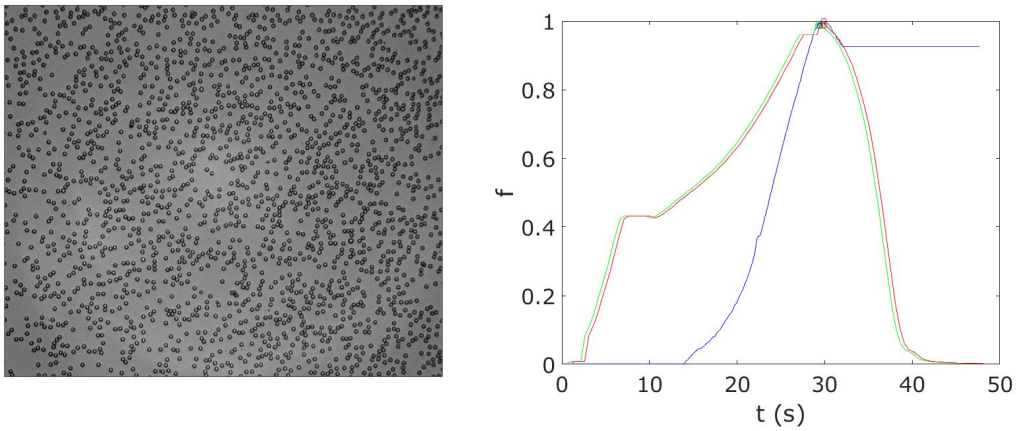
centrifugal microscopy can be used to measure bond lifetimes under force.

For centrifugal microscopy, a sample is placed in a tabletop centrifuge. The alignment of the samples is in such a way, that when the centrifuge rotates, the direction of the centrifugal force on the particles is away from the surface. The force applied to these particles can be calculated using:

$$F = 4\pi^2 f^2 mr. \tag{33}$$

with m the difference in mass between the particle and the solution, f the number of rotations per second and r the distance of the particles from the center of the centrifuge. Figure 24 illustrates the setup schematically [40].

During the acceleration and deceleration of the centrifuge, the sample shifts compared to the view of the camera. Since the acceleration of the sample causes this drift, the shift profile can be used to align the video with the force profile. Figure 25b shows an example of the agreement between the relative shift and relative velocity. I tracked the position of the particles and saved the times of the switches to find the transition times.



(a) *Experimental camera view.*

(b) *Alignment of normalized rotational velocity profile (green) and sample drift profile (blue).*

Figure 25: *(a) Example of an initial distribution of beads. Before the experiment starts, these particles move around in an area limited by the tethers. Upon acceleration, the tethers get stretched, and the movement of the particles is limited until the moment that a Nanoswitch opens. At that moment, the particle moves quickly to a new location. (b) Normalized rotational velocity profile (green) and sample drift (blue). The acceleration of the sample causes the sample drift, so the drift and acceleration profile can be aligned to connect lifetimes with dissociation forces. The red line shows the aligned acceleration profile.*

3.6 OxDNA: MD simulations of DNA

Simulations provide a different option to look at the behavior of catch bonds. An advantage of using simulations is that the position of individual bases can be tracked over time. For instance in simulations, we can track whether bases are bound or unbound. This information can be used for our DNA catch bonds to check which dissociation pathway is followed. Also, simulations can be repeated many times without the need for intervention of a researcher, making simulations less time-consuming.

Models for DNA range from the atomistic simulations (for example Ascona B-DNA Consortium [41]) including minute details of DNA, to continuum models, which can be used to simulate larger timescales ([42]). The disadvantage of detailed atomistic simulations is that within acceptable computation times, the simulated time is limited. The continuum models have the disadvantage that they cannot provide information about the hybridization and melting processes. Purely statistical models such as the previously described Nearest-Neighbor model and rate models provide limited information about the dynamics.

The working mechanism of the DNA catch bonds relies on the hybridization and melting of DNA, so enough detail is required to gain information about these processes. Furthermore, the simulated time should be long enough, since dissociation of the DNA catch bonds can take seconds. Due to the large variation in dissociation times and pathways, we need many dissociation times to find accurate average lifetimes and pathways. So the above-described methods cannot be used.

Several dynamical models allow for long simulations with more detail, by coarse-graining to the scale of nucleotides. However, these models often work only for specific conformations or do not include the structure of DNA in a realistic way [31]. A few models [31, 44] are exceptions, including structure using orientation dependent potentials. One of these models is the DNA specific Molecular Dynamics simulation program oxDNA developed in the J.P.K. Doye and A.A. Louis [31]. This program, when combined with Forward Stairway Sampling, allows for simulations with enough detail in a short enough time.

The used potentials in oxDNA are Finitely-extensible nonlinear elastic (FENE) springs for backbone interactions, Morse potentials for stacking interactions and H-bonding, harmonic potentials (for cross-stacking and coaxial stacking), Lennard-Jones potentials (for the soft repulsion of the backbone and bases), quadratic modulation potentials and quadratic smoothing potentials. These potentials are combined in such a way that excluded volume, stacking, hydrogen bonding, cross-stacking and coaxial stacking agree as much as possible with the results from experiments.

Not all aspects of DNA are represented in oxDNAs. First of all, hydrogen bonding besides A/T and G/C is not included, as well as sequence-specific interactions besides hydrogen bonding. This means this protocol does not include effects such as preferred sites of bubble formation. Also, the model does not explicitly include electrostatic interactions, making the model only realistic at high salt concentrations. A third approximation is the equal sizes of the DNA grooves. In reality, DNA has mayor and minor grooves, meaning that the backbone strands are closer together on one side of the helix than on the other side (in an alternating fashion). Finally, the potential used for the backbone is not very realistic, as bases can rotate and bend freely.

Although these approximations neglect important aspects of DNA, we do not expect that most of these approximations have a significant influence on our results. In our simulations, we use ideal DNA, in the sense that undesired hydrogen bonding is impossible and all bases

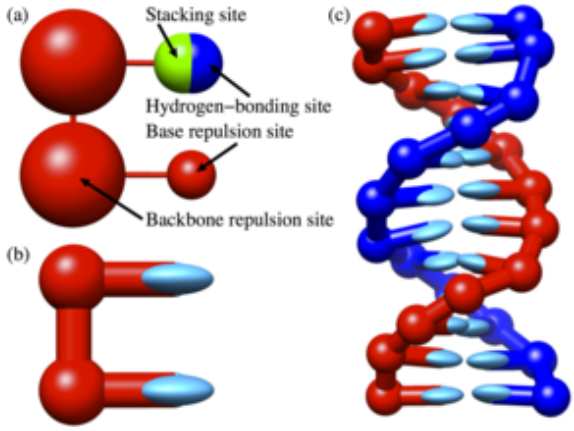


Figure 26: *(a) Interaction sites of oxDNA. Hydrogen-bonding and base repulsion interactions are shown on separate bases for clarity. The backbone is connected using a FENE spring potential. Hydrogen-bonding/stacking is represented by a combination of a Morse potential and a harmonic potential and repulsion by Lennard-Jones potentials. (b) Representation of bases in figures, not representative for interactions. (c) Example of a duplex. Figure from the oxDNA website [31]*

have the same strength. In this way, side effects caused by the used DNA sequence can be suppressed. So not including hydrogen bonding besides A/T and G/C bonding does not influence the results. The salt concentration in our simulations is high enough to neglect electrostatic interactions. The groove sizes are mainly crucial for protein binding and interactions between pieces of dsDNA [45], which are not tested in our simulations. The backbone potential is influential during our simulations because the propagation of forces through the backbone is a significant aspect of our simulations. So the unrealistic backbone potential could potentially make the simulations less realistic.

We use the Molecular Dynamics function of oxDNA, which uses the potentials to calculate the position of particles over time numerically. In this algorithm, the noise and damping are included using a random force, which is assumed equal for all translational and rotational modes. So the method does not include hydrodynamic effects. We relax the initial structure using Monte Carlo simulations with an additional linear potential bringing complementary bases close to each other.

OxDNA has been applied to describe various systems. These include structural studies of DNA such as the study of DNA Nano Tweezers [46], the interaction between hairpins [47] and DNA walkers [48]. Also, oxDNA has already been used to look at DNA under force. These studies include an analysis of the overstretching transition of DNA [49] and rupturing of DNA duplexes [50]. Appendix B describes the used simulation process for the results in this thesis.

3.7 Stairway Sampling: Reducing the computational time of simulations

Although oxDNA allows for relatively long and detailed simulations, the computational time is still too long to get accurate average dissociation times of DNA catch bonds. To solve this problem, I improved existing rare event sampling methods in such a way that they can be used to decrease the computational time to find accurate dissociation times. This method, called Stairway Sampling, is discussed in this section.

A common problem with brute-force simulations is that the real time you simulate in the available computational time is too short to get the required information. Luckily, there are many ways to decrease the needed computation time. Ways to influences this time can be classified into three levels. The required simulated real time can be decreased, the amount of real time within the simulated time can be increased and the amount of simulated time within the computation time can be increased. To decrease the needed computation time, you can improve your simulation method at each of these levels.

After optimizing your system, programming and (brute-force) simulation method, there are multiple common ways to improve the needed computation time further. Typical examples involve adding known virtual potentials driving the simulation toward the rare event, flattening the energy landscape or increasing the energy. These techniques decrease the needed real time. Coarse-graining your model is an example of a way to increase the simulation time within the computation time. This section describes a new method that increases the real time about which information is gathered within the simulated time. This type of methods can be used together with other methods to optimize the computational time, or without any of the other methods if their requirements and implications are undesired.

A group of already existing methods that are often used to increase the amount of real time within the simulated time is rare event methods. In simulations, rare events are events that are statistically very unlikely to happen within a given computation time through brute-force simulations. The goal of rare event sampling methods is reducing the computation time needed to get valuable information about rare events. Using these methods, transition rates can be estimated accurately while using less computational time.

Rare event sampling methods can also be used to find average transition times. Depending on how much real time can be simulated with the chosen simulation method, rare events can be as frequent as phase transitions of molecules or as rare as transitions of comets between solar systems [51]. Rare events are often not only described by the likelihood of occurrence, but also by the difference in associated timescales [52, 53]. A widespread assumption for the calculation of average transition times using Rare Event Sampling methods is that the total time of a rare event is negligible in comparison to the time scale associated with stable states.

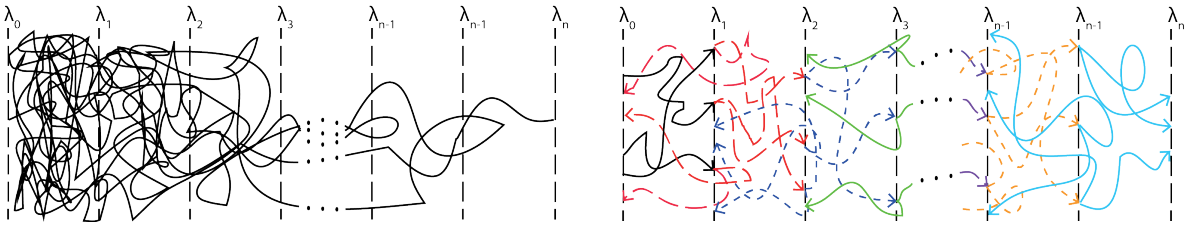
Rare Event methods can be divided into methods based on rate theory and path sampling [53, 54] or a combination of both. Rate theory generally uses knowledge of a barrier to estimate the reaction rates to and from neighboring states. For instance, for Transition State Theory to work, a transition state must be chosen at an optimal position in the energy landscape. Some examples are Transition State Theory, Reactive Flux Method, Kramers Rate Theory, Unimolecular Rate Theory, Mean-First-Passage-Time Approach and Quantum Rate Theory [55, 55, 56]. Some of these methods do not make any assumptions about the involved timescales, but all of these methods require knowledge of the shape of the energy landscape (barriers and initial and final states) or make assumptions about the shape of the energy landscape. As a result, this type of Rare Event methods can only be used in specific cases.

Path sampling methods reduce the computation time by reducing the sampling frequency of oversampled regions and increasing the sampling frequency of under-sampled regions. Some examples are Transition Path Sampling [55, 57, 58], Transition Interface Sampling (TIS) [55, 58–61], Replica Exchange TIS [55], Milestoning (MS) [62–64], Partial Path TIS (ppTIS) [61] and Forward Flux Sampling (FFS) [55, 65, 66]. Although some of these methods do not require knowledge of the shape of the energy landscape, for the calculation of the transition times they all require that the total rare event time is negligible in comparison to the timescale associated with stable states.

So for the calculation of average transition times of rare events, you need to make assumptions about the shape of the energy landscape or assumptions on the involved timescales. The new method Stairway Sampling eliminates these requirements, making it possible to calculate transition times of rare events in a larger number of situations. Because the requirement for the involved timescales disappears, the method is not only applicable as a Rare Event Sampling method but also to decrease the computation time to determine the transition time of other time-consuming events. Stairway Sampling can, for instance, be used for the simulation of relaxation of molecules to their ground-state, even if metastable states are passed along the way.

Stairway Sampling splits a simulation track in multiple short simulation tracks, defined by a reaction coordinate. Depending on the system under consideration, Stairway Sampling uses a full or a less strict Markovian assumption. Under a full Markovian approximation (Markovian approximation at every boundary), stairway sampling can decrease the computation time from increasing exponentially to increase linearly for a linear decreasing event probability. Reducing the Markovian approximation reduces the computational gain but makes the method wider applicable. Forward Stairway Sampling, based on Forward Flux Sampling, only uses a Markovian assumption at the starting point of the simulation. In this section, I show how Stairway Sampling defers from previous methods in its equations and applicability. Also, I discuss the reduction in computation time achieved with this method. As an example, I use Stairway Sampling for a simple Markovian chain process.

Rare event sampling methods With brute-force simulations, getting a sufficient amount of data about rare events takes much computation time, due to the unlikeliness of the rare event happening within the simulation time. Most of the time is spent simulating more com-



(a) *Example of a direct simulation path.* (b) *Example of a set of path sampling tracks.*

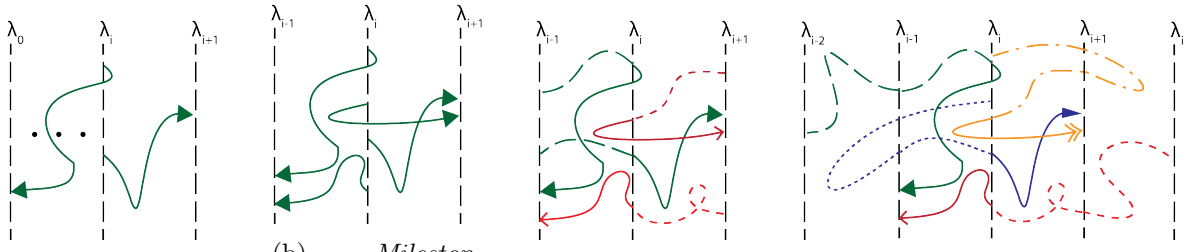
Figure 27: *Schematic representation of an example of a direct simulation track and a combination of path sampling tracks. In the direct simulation, most of the simulation time is spent in a stable region at the left, and not much data is gathered about the region on the right. With the example of the path sampling method, there is an even distribution of simulation time for all regions. Individual tracks are combined to find information about the full tracks.*

mon events. At the moment you start gathering the first bits of data for the rare events, you already have more than enough information about common events or tracks within stable regions. Rare event sampling uses this insight to save computation time.

There are already multiple rare event sampling methods to calculate transition times in such a way that it is unnecessary to have prior knowledge about the energy landscape. Three examples are Forward Flux Sampling (FFS), Milestoning (MS) and partial path Transition Interface Sampling (ppTIS). In these methods, a simulation is divided into multiple short simulations. The information from these simulations can be combined to find transition rates, times and probabilities, which subsequently can be used to find other information about the simulated system.

To split the tracks in FFS, MS and ppTIS, an order parameter must be defined. This order parameter separates phase space in such a way that a full track must visit all values of this order-parameter. Figure 27 shows an example of such a division in states. To be able to calculate the transition times using one of these methods, you must assume that the time spent in the initial state is long in comparison to the time spent in other states. When applying these methods, strict restrictions must be followed for the initial state to get correct results [54]. These limitations also make these methods unsuitable for the transition time calculations of anything else than rare events.

Forward Flux Sampling (FFS) In Forward Flux Sampling, phase space is divided into states by a series of boundaries that must be crossed to go from the initial state to the final state. Instead of starting each simulation in the initial state and waiting until the final state is reached, new simulations are started at the boundaries between the states and continued until the next state or the first state is reached. In the first state, the transition times are saved and in the other states the probabilities to move to the next state (instead of to the first state). Figure 28 shows a schematic of possible tracks from one boundary for Forward Flux Sampling.



(a) *FFS/FSwS* (b) *Milestoning/MSwS* (c) *ppTIS/DSwS* (d) *2th order SwS*

Figure 28: *Tracks assumed different in simulation methods. With direct simulations, all tracks are simulated from the start till the end, not assuming any interchangeability of parts of tracks. Forward Flux sampling assumes tracks at the first boundary are independent and can be combined to form complete tracks (Markovian at first boundary). Milestoning assumes all states are independent and tracks from state to state can be combined to form complete tracks (0^{th} order Markovian). Partial Path Transition Interface Sampling assumes tracks, given their previous boundary, can be interchanged to form new tracks (1^{st} order Markovian). Higher order approximations can be made with the same logic. In the figure, color, arrowheads and the dashing of the history line show which lines are considered different in each method.*

Forward Flux Sampling relies on the assumption that the distribution at the starting boundary does not change for tracks going to a higher boundary and back. This First State Markovian approximation means that the first two boundaries should be separated enough, so the distribution of tracks at these boundaries is uncorrelated. For transition time calculations, Forward Flux Sampling also assumes that the number of tracks in the first state is approximately equal to the total number of tracks ($\frac{N_F}{N_0} \approx 1$), which means that the tracks at the first state are much longer than the other tracks ($t_0 \gg t_i$) or that the tracks have a negligible probability to be at another state ($P_0 \gg 1$).

That Forward Flux Sampling relies on a timescale assumption becomes clear from the derivation of the transition time equations. The flux from initial state 0 to final state n is given by

$$\begin{aligned} \Theta_{0\to n} &= \Theta_{\overrightarrow{0\to n}} - \Theta_{\overleftarrow{0\to n}}; \\ &= N_F k_{0\to n} - N_n k_{0\leftarrow n}. \end{aligned} \tag{34}$$

with $k_{i\to j}$ the rate from state i to state j and N_F the number of forward tracks. The average rate is the inverse of the average first passage time ([67]). So:

$$\Theta_{\overrightarrow{0\to n}} = N_F k_{0\to n} = \frac{N_F}{\tau_{0\to n}}, \tag{35}$$

with $\tau_{i\to j}$ the average time from state i to state j.

Now let us look at the derivation of the same flux from the splitted tracks in Forward Flux Sampling. First, the flux to the first boundary (of all tracks in state 0, N_0) can be calculated in a similar way as the total flux:

$$\Theta_{\overrightarrow{0\to 1}} = \frac{N_0}{\tau_{0\to 1}}. \tag{36}$$

The number of tracks reaching the second boundary is equal to the number of tracks reaching the first boundary, times the fraction of these tracks going on to the second boundary. Over time, this is equal to

$$\Theta_{\overrightarrow{0\to 2}} = \frac{N_0}{\tau_{0\to 1}} P_{1\to 2}. \tag{37}$$

In this equation, $P_{i\to j}$ is the fraction of constructs in state i going to state j. This process can be repeated for each following boundary, resulting in

$$\Theta_{\overrightarrow{0\to n}} = \frac{N_0}{\tau_{0\to 1}} \prod_{i=0}^{n-1} P_{i\to i+1} \tag{38}$$

for n boundaries. So given the number of events in the first state (N_0), the transition probabilities ($P_{i\to i+1}$) and the first transition time, an estimation can be made for the average flux from the starting point to the endpoint.

To find the transition times, equation 35 and equation 38 are combined:

$$\tau_{0\to n} = \frac{N_F}{N_0} \frac{\tau_{0\to 1}}{\prod_{i=0}^{n-1} P_{i\to i+1}}. \tag{39}$$

Forward Flux Sampling only saves the transition times at the first state. This means that we can only draw conclusions about the transition times if the transition time in the first state

is the only significant transition time. In other words, $t_0 \gg t_i$ for all i or $P_0 \approx 1$. These assumptions are equivalent to assuming $N_F \approx N_0$. So the final transition time equation for forward flux sampling is:

$$\tau_{0\to n}^{FFS} = \frac{\tau_{0\to 1}}{\prod_{i=0}^{n-1} P_{i\to i+1}} = \frac{\tau_{0\to 1}}{P_{0\to n}}. \tag{40}$$

So when you use Forward Flux Sampling to calculate transition times, you need to make assumptions on the involved timescales.

Milestoning (MS) Milestoning is a similar technique to Forward Flux Sampling, but instead of continuing simulations to the next or first state, simulations are continued to the next or previous state. This drastically reduces the simulation time but also tightens the Markovian approximation. Not only the first boundary, but all boundaries must be Markovian for this method to be applicable.

The resulting transition time equation for Milestoning is:

$$\tau_{0\to n}^{MS} = \tau_{0\to 1} \sum_{j=1}^{n} \prod_{k=1}^{j-1} \frac{p(\substack{k\to k-1\\ k+1})}{p(\substack{k\to k+1\\ k-1})}. \tag{41}$$

The probabilities ($p(\substack{k\to k+1\\ k-1})$) are the chance for a track to move one step forward, instead of backward (and vice versa). These probabilities differ from the probabilities used in Forward Flux Sampling. These probabilities are, however, closely related to each other, revealing the assumptions made in Milestoning.

The probabilities used in Forward Flux Sampling can be expressed in the probabilities used in Milestoning. To find the corresponding equation, we express the probability $P_{0\to i}$ in $P_{0\to i-1}$ and the Milestoning probabilities. Using that $p(\substack{i\to i+1\\ i-1}) + p(\substack{i\to i-1\\ i+1}) = 1$ and solving for $P_{0\to i}$ results in [61]:

$$\begin{aligned} P_{0\to i} &= \frac{p(\substack{i-1\to i\\ i-2})P_{0\to i-1}}{p(\substack{i-1\to i\\ i-2}) + p(\substack{i-1\to i-2\\ i})p(\substack{i-2\to 0\\ i-1})}; \\ p(\substack{i\to 0\\ i+1}) &= \frac{p(\substack{i\to i-1\\ i+1})p(\substack{i-1\to 0\\ i})}{p(\substack{i\to i+1\\ i-1}) + p(\substack{i\to i-1\\ i+1})p(\substack{i-1\to 0\\ i})}. \end{aligned} \tag{42}$$

These equations can be used in the transition time equation for Forward Flux Sampling and solved recursively. This results in the transition time equation for Milestoning, showing that Milestoning uses the same assumption as Forward Flux Sampling.

Partial Path Transition Interface Sampling (ppTIS) Partial path Transition Interface Sampling (ppTIS) also simulates tracks till the previous or next state but differentiates between tracks coming from the previous state and next state. This reduces the full (0^{th} order) Markovian approximation of Milestoning to a 1^{st} order Markovian approximation.

The transition time of ppTIS is given by:

$$\tau_{0\to n}^{ppTIS} = \tau_{0\to 1} \sum_{j=0}^{n-1} \frac{p(\substack{j\to j-1\\ +\quad j+1})}{p(\substack{j\to j+1\\ +\quad j-1})} \prod_{k=1}^{j-1} \frac{p(\substack{k\to k-1\\ -\quad k+1})}{p(\substack{k\to k+1\\ +\quad k-1})}. \tag{43}$$

for ppTIS. In these equations, $p(^{i\to j}_{k})$ is the probability to go from state i tot state j instead of k. $p(^{i\to j}_{+\ k})$ and $p(^{i\to j}_{-\ k})$ are the probabilities for tracks passing boundary (i) going forward $(+$, coming from $i-1)$ or backward $(-$, coming from $i+1)$.

Similar as for Milestoning, the probabilities of Forward Flux Sampling can be expressed in the probabilities of ppTIS. This results in [61]:

$$\begin{aligned} P_{0\to i} &= \frac{p(^{i-1\to i}_{+\quad i-2})P_{0\to i-1}}{p(^{i-1\to i}_{+\quad i-2}) + p(^{i-1\to i-2}_{+\quad\ i})p(^{i-2\to 0}_{-\quad i-1})}. \\ p(^{i\to 0}_{-\ i+1}) &= \frac{p(^{i\to i-1}_{-\ i+1})p(^{i-1\to 0}_{-\ i})}{p(^{i\to i+1}_{+\ i-1}) + p(^{i\to i-1}_{+\ i+1})p(^{i-1\to 0}_{-\ i})}. \end{aligned} \tag{44}$$

Using these equations in the transition time equation for Forward Flux Sampling (and solving recursively) results in the transition time equation for ppTIS. So, like FFS and MS, ppTIS relies for its transition time calculation on the assumption that the time spent in all other states than the first state is negligible in comparison to the time spent in the first state..

Other Rare Event Sampling Methods Similar Rare Event Sampling Methods can be constructed with even higher order Markovian approximations. Although derivations get more cumbersome, the idea that Forward Flux Sampling probabilities can be expressed in the methods probabilities upholds. This means that all methods based on this logic require $t_0 \gg t_i$ or $P_0 \approx 1$.

All methods that only save the transition probabilities and not the transition times to and from each state must make assumptions on the involved timescales. So these methods not applicable in every situation.

Stairway Sampling Stairway Sampling uses the path sampling schemes above but does not assume that approximately all time is spent in the first state. Like for the above-described Path Sampling methods, the method can be used with a full Markovian approximation (and maximal speedup) or a less strict Markovian approximation and a lower speedup. For the calculation of the transition time, the protocol needs not only the transition time between the first boundaries and the transition probabilities, but also the transition times in other states.

Forward Stairway Sampling (FSwS) Forward Stairway Sampling is the Stairway Sampling method based on Forward Flux Sampling, uses a Markovian approximation at the first boundary. There are two ways to find the transition time equations for this method. First, the ratio $\frac{N_F}{N_0}$ can be solved from the involved differential equations for equilibrium and steady-state situations. Second, the transition time can be found directly by combining transition times and probabilities by recursively solving equations. This direct derivation shows that the method can be applied out of equilibrium and without constant state occupation.

Let us first connect our theory to the previously derived theory of Rare Event Sampling methods. We saw before that the transition times without assumptions on the involved timescales is given by 39:

$$\tau_{0\to n} = \frac{N_F}{N_0}\frac{\tau_{0\to 1}}{\prod_{i=0}^{n-1} P_{i\to i+1}}. \tag{45}$$

So all we need to do is find the ratio $\frac{N_F}{N_0}$. This ratio can be found by solving the involved differential equations:

$$\frac{\mathrm{d}N_i}{\mathrm{d}t} = \Theta_{i-1\to i} - \Theta_{i\to i+1} - \Theta_{i\to 0}. \tag{46}$$

with $\Theta_{j\to k}$ the flux from state j to state k.

In the case of equilibrium or a steady state, at every moment in time $\frac{N_i}{\tau_i} = \frac{N_i}{P_{i\to i+1}\tau_{i\to i+1}+P_{i\to 0}\tau_{i\to 0}}$ tracks arrive (and leave) state i. Probabilities to move in forward or backward determine the size of the flux in each of the directions:

$$\begin{aligned}
\Theta_{i-1\to i} &= \frac{N_{i-1}}{\tau_{i-1}} P_{i-1\to i};\\
\Theta_{i\to i+1} &= \frac{N_i}{\tau_i} P_{i\to i+1};\\
\Theta_{n\to 0} &= \frac{N_i}{\tau_n} P_{n\to 0}.
\end{aligned} \tag{47}$$

So equation 46 can also be written as:

$$\frac{\mathrm{d}N_i}{\mathrm{d}t} = \frac{N_{i-1}P_{i-1\to i}}{\tau_{i-1}} - \frac{N_i P_{i\to i+1}}{\tau_i} - \frac{N_i P_{i\to i+1}}{\tau_i} = 0 \tag{48}$$

Solving this equation for N_i and using the result iteratively gives:

$$N_n = N_0 \frac{\tau_n}{\tau_{0\to 1}} \prod_{i=0}^{n-1} P_{i\to i+1}; \tag{49}$$

This result physically means that the number of tracks at a boundary is equal to that part of the number of tracks starting at λ_0 and reaching the boundary multiplied by the time they spent at that boundary relative to the time they spent at λ_0.

The total number of tracks is equal to $\sum_{i=0}^{n} N_i = N_F + N_n$. To get $\frac{N_F}{N_0}$ we need to add the solutions for all boundaries except the last boundary λ_n (all forward tracks) and divide by the number of tracks in the first state. Combining the result with equation 39 results in

$$\tau_{0\to n} = \sum_{j=0}^{n-1} \frac{\tau_j}{\prod_{i=j}^{n-1} P_{i\to i+1}}, \tag{50}$$

the transition time equation for Forward Stairway Sampling.

The same result can be derived using just a combination of rate equations. To use Stairway Sampling, an order parameter must be defined, dividing space into states. In Forward Stairway Sampling, these states are not independent. So each state λ_k, consists of substates $\lambda_{k,i}$ that each behave very differently. For Forward Stairway Sampling to be applicable in every situation, all these states must be taken into account and not change the final result.

We start with the definition of the average probabilities and times in a system with m_k substates in state k and m_l substates in state l. The probability to move from state k to state l is the sum of all probabilities to move from one of the substates of state k to state l, weighted by the probability to be in each of these states. The average time of a transition from state k to l is the weighted average of the average transition time of the individual substates. Not all tracks are from k to l, hence the weighting must be corrected with a factor $\frac{1}{P_{k\to l}}$. So:

$$
\begin{aligned}
P_{k\to l} &= \frac{1}{\sum_{i=1}^{m_k} P_{k,i}} \sum_{i=1}^{m_k} P_{k,i} \sum_{j=1}^{m_l} P_{k\to l,i\to j} \\
\tau_{k\to l} &= \frac{1}{P_{k\to l} \sum_{i=1}^{m_k} P_{k,i}} \sum_{i=1}^{m_l} P_{k,i} \sum_{i=1}^{m_l} P_{k\to l,i\to j} \tau_{k\to l,i\to j};
\end{aligned}
\tag{51}
$$

Now we express the transition time from an initial state λ_0 to state λ_n in the transition time from state λ_0 to state λ_{n-1} and from state λ_{n-1} to λ_n. This can either be directly, or after first returning (multiple times) to λ_0. However, because the tracks are assumed independent at the first state, the time after returning to the initial state is equal to $\tau_{0\to n}$. So:

$$
\begin{aligned}
\tau_{0\to n} = &\frac{1}{\sum_{i=1}^{m_0} P_{0,i}} \sum_{i=1}^{m_0} P_{0,i} \sum_{i=1}^{m_{n-1}} P_{0\to n-1,i\to j} \tau_{0\to n-1,i\to j} \\
&+ \frac{1}{\sum_{i=1}^{m_{n-1}} P_{n-1,i}} \sum_{i=1}^{m_{n-1}} P_{n-1,i} \sum_{i=1}^{m_n} P_{n-1\to n,i\to j} \tau_{n-1\to n,i\to j} \\
&+ \frac{1}{\sum_{i=1}^{m_{n-1}} P_{n-1,i}} \sum_{i=1}^{m_{n-1}} P_{n-1,i} \sum_{i=1}^{m_0} P_{n-1\to 0,i\to j} \left(\tau_{n-1\to 0,i\to j} + \tau_{0\to n}\right)
\end{aligned}
\tag{52}
$$

Using the definitions of average times and probabilities:

$$
\begin{aligned}
\tau_{0\to n} = &P_{0\to n-1} \tau_{0\to n-1} + P_{n-1\to n} \tau_{n-1\to n} \\
&+ P_{n-1\to 0} \left(\tau_{n-1\to 0} + \tau_{0\to n}\right).
\end{aligned}
\tag{53}
$$

Solving for $\tau_{0\to n}$ results in an expression of $\tau_{0\to n}$ in $\tau_{0\to n-1}$. Replacing n with $n-1$, $n-2$, enz. we can use the same reasoning at all states. Combining this information gives the same result as found before for equilibrium and steady-state problems:

$$
\tau_{0\to n}^{FSwS} = \sum_{i=0}^{n-1} \frac{\tau_i}{\prod_{j=i}^{n-1} P_{j\to j+1}}.
\tag{54}
$$

This shows that Forward Stairway Sampling is applicable in other situations than rare events. By taking the limits $P_0 \gg P_i$ or $\tau_0 \gg \tau_i$ the equation reduces to the equation for Forward Flux Sampling. Using $P_i = 1$ gives the expected result of a total time equal to the sum of forward times.

Forward Stairway Sampling (and Forward Flux Sampling) use even less computational time when tracks are certain are pruned. Most of the FSwS computation time is spent simulating tracks moving from a boundary back to the initial state. If more statistics than necessary is generated for these transitions, part of the tracks can be truncated at the previous boundary., resulting in an additional decrease of the computation time.

Markovian Stairway Sampling (MSwS) Markovian Stairway Sampling is the 0^{th} order Markovian version of Stairway Sampling. Tracks are continued until they reach the next or previous state. In the process of measuring the transition times and probabilities, you start with tracks starting at the first state, then the tracks starting in the second state and so on until the last state. So backward tracks are never used, so states must by fully Markovian.

The equation for the transition times at each interface can be found found by addition of individual transition times. Since states are assumed independent, we do not have to take substates into account and can immediately use average times and probabilities at states. To simplify the derivation, we define $T^{i\to i+1}$ as the average extra time needed to reach boundary $i+1$, following every possible path. So the average time to reach $i+1$ from λ_0 is given by $\tau_{0\to i+1} = \tau_{0\to i} + T^{i\to i+1}$ and in general:

$$\tau_{0\to n} = \sum_{i=1}^{n} T_{i-1\to i}. \tag{55}$$

Now we will express $T_{i-1\to i}$ in individual average times and probabilities. Due to the Markovian approximation, we can say that:

$$\begin{aligned} T^{i-1\to i} &= p(\substack{i-1\to i\\ i-2})\tau(\substack{i-1\to i\\ i-2}) \\ &+ \quad p(\substack{i-1\to i-2\\ i})\left(\tau(\substack{i-1\to i-2\\ i}) + T^{i-2\to i-1} + T^{i-1\to i}\right) \end{aligned} \tag{56}$$

Solving for $T_{i-1\to i}$ gives:

$$T^{i-1\to i} = \frac{\tau^{i-1} + p(\substack{i-1\to i-2\\ i})T^{i-2\to i-1}}{p(\substack{i-1\to i\\ i-2})} \tag{57}$$

with τ^i the average time from boundary i. This equation can be solved recursively, giving:

$$\tau_{0\to n}^{MSwS} = \sum_{i=1}^{n}\sum_{j=0}^{i-1} \frac{\tau^j}{p(\substack{j\to j+1\\ j-1})} \prod_{k=j+1}^{i-1} \frac{p(\substack{k\to k-1\\ k+1})}{p(\substack{k\to k+1\\ k-1})}. \tag{58}$$

Taking the limit $t_0 \gg t_i$ or $P_0 \gg P_i$ reduces the equation to the equation for Milestoning. Setting the forward probabilities to 1 again results in a total time equal to the summation of forward times.

Directional stairway sampling (DSwS) Directional Stairway Sampling is the 1st order Markovian version of Stairway Sampling. This method differentiates between tracks coming from the previous state and the next state. To get all transition times and probabilities, simulations are started from the initial state, followed by simulations at higher states where the final states of tracks from the previous state are used as starting points. This process must be repeated in the opposite direction, from the final state to the first state, to get information about tracks starting in the opposite direction.

Derivation, similar to the derivations of FSwS and MSwS, results in the following equation:

$$\begin{aligned} \tau_{0\to n}^{DSwS} = \sum_{i=0}^{n-1} \Bigg[& \frac{\tau_+^i + p(\substack{i\to i-1\\ +\quad i+1})\tau_-^{i-1}}{p(\substack{i\to i+1\\ +\quad i-1})} \\ & + \frac{p(\substack{i\to i-1\\ +\quad i+1})}{p(\substack{i\to i+1\\ +\quad i-1})} \sum_{j=1}^{i-1} \left(\tau_+^j + \tau_-^{j-1}\right) \prod_{k=j}^{i-1} \frac{p(\substack{k\to k-1\\ -\quad k+1})}{p(\substack{k\to k+1\\ +\quad k-1})} \Bigg]. \end{aligned} \tag{59}$$

If the forward and backward probabilities and times are equal, the system is again fully Markovian, and the equations reduce to the equations for MSwS. Assuming $t_0 \gg t_i$ or $P_0 \gg P_i$, the equation reduces to the equation for ppTIS. Setting all forward probabilities to 1 again results in a total time equal to the summation of forward times.

Higher order Stairway Sampling The applicability of the method can be improved further by increasing the amount of history included in each track. Markovian Stairway Sampling (0th order Markovian) has the highest speed-up but requires fully independent states. Directional Stairway Sampling (1st order Markovian) reduces the speed-up slightly but improves the applicability. You can use even higher order Markovian approximations to make the method even wider applicable. 2^{rd} Order Markovian Stairway Sampling uses the previous two states of each track for differentiation, 3^{rd} Order Markovian Stairway Sampling the previous three boundaries. One can continue in this way, improving the applicability and decreasing the computational gain, until the full history of tracks is used.

The schemes discussed so far do not use any knowledge about the system under consideration. Sometimes, a higher accuracy or speed-up can be achieved by choosing multiple order parameters. In that way, a distinction can be made between tracks, for instance, because there are multiple distinct pathways. Also, it might be valuable to combine the different Stairway Sampling schemes to get the maximal speed-up and accuracy. For instance, the total transition times of subtracts found using Forward Stairway Sampling can be combined using Markovian Stairway Sampling to get information about the full system. So although knowledge about the explored system is not necessary, it can be used to improve your computational time and accuracy.

Applicability and speed-up Stairway Sampling can be applied in many situations. Here we apply it to a simple Markovian Chain model. This shows how Stairway Sampling can be used to decrease computational time, even when Rare Event Sampling methods are not applicable.

Applicability Stairway Sampling Stairway Sampling can be applied to decrease computation time in all situations where regions in phase space are oversampled in comparison to other regions. To show that Stairway Sampling is wider applicable than existing Rare Event Sampling methods, we compare the different versions of Stairway Sampling with Forward Flux Sampling, Milestoning and Partial Path Transition Interface Sampling.

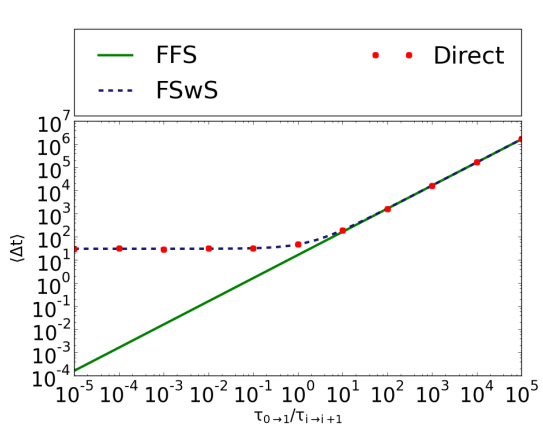
We define a simple process consisting of 5 states. In each time step, a track either moves one step forward or back to the start. We can tune the probabilities and time per time step within this process. The process finishes when the fifth state is reached. This straightforward simulation scheme can be solved analytically and, even with direct simulations, the simulation time is small. So also direct, brute-force simulations are possible. To test Markovian Stairway Sampling and Directional Stairway Sampling, we use a similar scheme, but now the track either moves one state forward or moves one state backwards. Also, we make it possible to alter the probabilities for tracks coming from the next state and the previous state.

Figure 29 shows the results of the direct simulations and the analytical results of the different Stairway Sampling methods and Rare Event Sampling Methods. When P is small, so $P_0 \gg P_i$ and $\tau_{0\to1} \gg \tau_{i\to j}$, all methods agree well with the direct simulations. For lower values of P or $\tau_{0\to1}$, only Stairway Sampling agrees with direct simulations. If the transition probabilities (or times) of tracks moving forward and backwards through a boundary are different, only Directional Stairway Sampling agrees over the full range. Of course, higher order differences could be introduced, resulting in only higher order versions of Stairway Sampling agreeing with direct simulations. So to conclude, Stairway Sampling gives results agreeing with direct simulations in a wider range of situations than the other discussed Rare

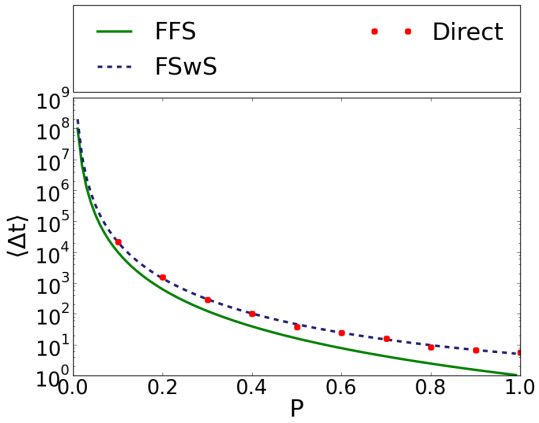
Event Sampling methods.

Difference computational time Besides the applicability, the reduction in computational time is the most important characteristic of Stairway Sampling. The reduction of computational time can be predicted analytically using the transition time equations.

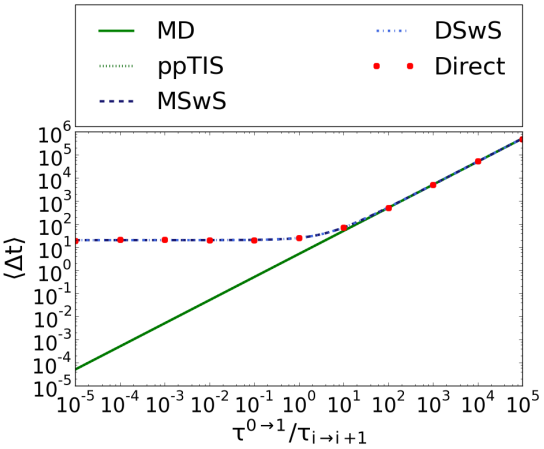
The computational time needed to simulate a process directly can be approximated by multiplying the number of time steps in the full process by the average computational time of one time step ($\tau_{0\to n} * \Delta t_{comp}/\Delta t_{step}$). This time can be expressed in the times of individual simulations using stairway sampling using the transition time equations (54, 58, and 59) for



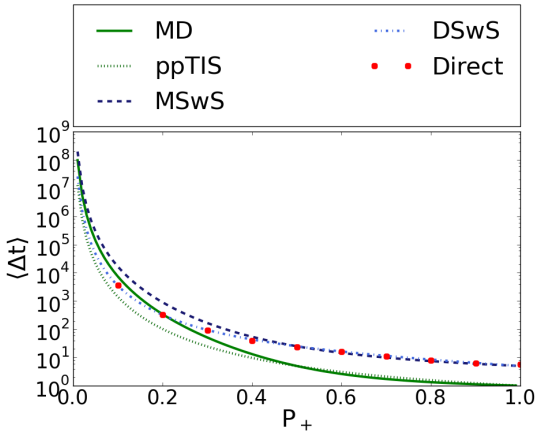
(a) *Dependence of the the total transition time on the relative time spent at the first state.*



(b) *Transition time dependence on the forward transition probability.*



(c) *Dependence of the the total transition time on the time at the first boundary, for a system where times and probabilities for forward and backward crossings are equal.*



(d) *Transition time dependence on the forward transition probability.*

Figure 29: *Dependence of the the total transition time forward transition times and backward transition times. Times are set to 1 and probabilities to 0.5, unless indicated differently.*

$\tau_{0\to n}$.

Using direct simulations regions are over-sampled, and other regions are under-sampled. The most under-sampled region is the bottleneck, causing the highest uncertainty. If we split the simulation space from the initial to the final state using an order parameter, there is always minimally one region with only one forward transition between defined intermediate states. This region is the least sampled regions, having the highest contribution to the uncertainty. So we can approximate the needed amount of data to get equally valuable statistics by the average time needed to get one forward track in every state.

In each of the Stairway Sampling methods, on average one in every $1/p$ tracks is forward. p is here the average forward probability, defined by $P_{i\to i+1}$ and $p(^{i\to i+1}_{i-1}$ for FSwS and MSwS. So for these methods, we need to simulate $1/p$ tracks in every state to get statistics comparable to one direct simulation. The resulting average needed computation time is

$$\begin{aligned} t_{comp}^{FSwS} &= \sum_{i=0}^{n-1} \frac{\tau_i}{P_{i\to i+1}} \frac{\Delta t_{comp}}{\Delta t_{step}}; \\ t_{comp}^{MSwS} &= \sum_{i=0}^{n-1} \frac{\tau^i}{p(^{i\to i+1}_{i-1})} \frac{\Delta t_{comp}}{\Delta t_{step}}. \end{aligned} \tag{60}$$

For DSwS we need both a transition with a forward and backward passing history, so:

$$t_{comp}^{DSwS} = \sum_{i=0}^{n-1} \left[\frac{\tau_+^i}{p(^{i\to i+1}_{+\ i-1})} + \frac{\tau_-^i}{p(^{i\to i+1}_{-\ i-1})} \right] \frac{\Delta t_{comp}}{\Delta t_{step}}. \tag{61}$$

Taking the ratio of the SwS simulation time and the direct simulation time shows how much the simulation time is reduced with each of the methods:

$$\begin{aligned} \Theta_{comp}^{FSwS} &= \frac{\sum_{i=0}^{n-1} \frac{\tau_i}{P_{i\to i+1}}}{\tau_{0\to n}^{FSwS}}; \\ \Theta_{comp}^{MSwS} &= \frac{\sum_{i=0}^{n-1} \frac{\tau^i}{p(^{i\to i+1}_{i-1})}}{\tau_{0\to n}^{MSwS}}; \\ \Theta_{comp}^{DSwS} &= \frac{\sum_{i=0}^{n-1} \left[\frac{\tau_+^i}{p(^{i\to i+1}_{+\ i-1})} + \frac{\tau_-^i}{p(^{i\to i+1}_{-\ i-1})} \right]}{\tau_{0\to n}^{DSwS}}. \end{aligned} \tag{62}$$

If we use $n = 1$ in these equations (remembering that the forward probability at the first barrier is 1), we find $\Theta_{comp}^{FSwS} = \Theta_{comp}^{MSwS} = \Theta_{comp}^{DSwS} = 1$. As expected, the computational times are equal to the computational times of direct simulations.

The computation time ratios become more instructive if we assume all probabilities and transition times are equal, $n > 4$ and that $p \ll 1$ (except the probability to move from the first to the second state, which is always 1). This corresponds to a simulation over a steep and always increasing barrier. Simplifying Θ results in:

$$\begin{aligned}
\Theta_{comp}^{FSwS}\begin{pmatrix} P_{i\to i+1} = P \ll 1 \\ \tau_i = \tau \end{pmatrix} &= \frac{n}{2}P^{n-2}; \\
\Theta_{comp}^{MSwS}\begin{pmatrix} p(^{i\to i+1}_{\ \ i-1}) = p \ll 1 \\ \tau^i = \tau \end{pmatrix} &= \frac{n}{2}p^{n-2}; \\
\Theta_{comp}^{DSwS}\begin{pmatrix} p(^{\ i\to i+1}_{-\ \ i-1}) = p \ll 1 \\ \tau_+^i = \tau_-^i = \tau \end{pmatrix} &= np^{n-2}.
\end{aligned} \tag{63}$$

These results show that DSwS is a factor two slower than MSwS. Higher order approximations are a factor 2^d slower, with d the order of the Markovian approximation.

The speedup using MSwS is higher than the speedup of FSwS because the forward probability P of FSwS is higher than the forward probability p of MSwS. The forward probability of MSwS is the probability to move one state forward instead of backward. The forward probability of FSwS is the probability to move one state forward instead of moving back to the first state. This probability is the probability to move forward directly (the forward probability of MSwS) plus the probability to first move back to one of the previous states and then move to the next boundary. For each system, there is an optimal number of states. Adding a boundary increases the exponent $n-2$, decreasing the needed time, but also increases the forward probability. The optimal choice of boundaries is the number of barriers where Θ is minimal.

In the same way as for Forward Flux Sampling, the computational time of Forward Stairway Sampling can decrease further by pruning backward tracks. On average 1 in every $1/p$ tracks moves forward, and all other tracks move backwards. So backward tracks are oversampled in comparison to forward tracks. By stopping part of these backward tracks when they reach the previous boundary and increasing the weight of the continued tracks, extra computational time can be saved.

Again using the requirement that each transition area is sampled at least once for each direct simulation, gives that 1 in every $1/p$ tracks can be continued. Most of the computational time is spent simulating these backward tracks since they are on average much longer than the forward tracks. So pruning reduces the computational time maximally by a factor $1/p$, giving:

$$\Theta_{comp}^{FSwS}\begin{pmatrix} P_{i\to i+1} = P \ll 1 \\ \tau_i = \tau \end{pmatrix} = \frac{n}{2}P^{n-1}. \tag{64}$$

In reality, Stairway Sampling speed-ups simulations less than above mentioned theoretical maxima. Extra time is spent on data management, such as loading and saving intermediate states, saving transition times and probabilities and randomly selecting states. Also, an additional systematic error could emerge for correlated states that are assumed to be independent. The lower sampling of regions that are oversampled in direct simulations also increases the uncertainty of Stairway Sampling.

4 Results: Testing DNA catch bonds

The theoretical section demonstrated that in theory DNA can be used to create a humanmade, tunable catch bond. In theory, all of the following requirements are satisfied:

- Stability and correct formation;
- Force induced state-switching;
- Hybridization of the extra, stabilizing duplex;
- Increase of lifetime after the state-switching and formation of the extra duplex;

Also, the lifetime-force dependency has a range of forces where the lifetime increases:

- Lifetime-force dependency.

In this chapter, I show whether these in theory satisfied requirements are also satisfied in our experiments and simulations. Also, I discuss the results three used ways of measuring the lifetime-force dependency.

The functionality of DNA catch bonds was tested in two ways: By confirming that each of the requirements is satisfied and by the functionality of the catch bond as a whole. To confirm that all requirements are satisfied, we tested the stability of a real DNA catch bond construct using gel electrophoresis in combination with restriction enzymes (section 4.1.1). Next, state-switching is tested experimentally using both force rate experiments and constant force experiments (section 4.1.2). The results of the optical tweezer experiments also indicate whether the extra, stabilizing region forms (section 4.1.3). Finally, the results of estimating the difference in lifetime between the folded and the extended state are discussed in section 4.1.4. The functionality of the DNA catch bonds as a whole was tested by measuring the lifetime-force dependency. Section 4.2 discusses the experimental and simulational results.

4.1 Testing DNA catch bond requirements

4.1.1 Stability

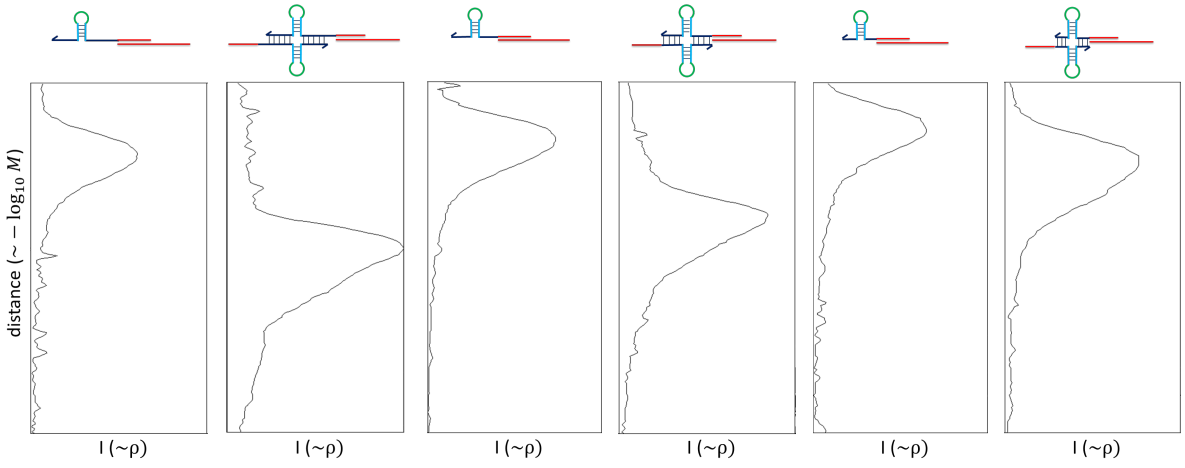
The DNA catch bonds only work, when the initial state of the DNA construct is the folded state. We check this in three steps. First, we check whether the two DNA strands involved in the construct, strand X and strand Y, indeed form a construct. By increasing the temperature, we also check whether the extended state can be formed, by checking whether the B/B'-region forms.

The two sides of a DNA catch bonds should bind. Whether this is happening in reality, is checked by performing gel electrophoresis experiments for three different strand sizes (N_H = 7,N_H = 10 and N_H = 25). For each of these constructs, I made two solutions, one with only strand X and one with both strand X and strand Y (and in both cases an attachment strand of 120 bases). Using each of the resulting solutions in one electrophoretic lane gives the results as shown in figure 30. At 4°C, all variations in the length of N_H show a difference in distance traveled between strand X and strand X and Y. Also, there is no peak observable

at the distance traveled by only strand X in the lanes with X and Y, suggesting all of the strands are at least part of the time bound to other strands.

The smaller constructs ($N_H = 7$) travel further than the large constructs ($N_H = 25$), as expected. Also, the difference between the distance traveled by strand X and Y and only strand X is more substantial for the bigger constructs. This difference is as expected since the relative difference in length is also smaller due to the attached strand of 120 bases. Comparing the distance traveled by strand X of $N_H = 25$ (262 bases) with a DNA ladder shows that the strands end between the ladder strands corresponding to 200 and 300 bases (100 and 150 base pairs). Comparing the distance traveled by both strands of $N_H = 25$ (404 bases) with a DNA ladder shows that the strand ends at approximately the same distance in the gel as ladder strand of 400 bases (200 base pairs). So the distances traveled in this experiment agree well with what is expected based on DNA ladders. Gel pictures can be checked in appendix C.

In the second set of experiments, I check the formation of the hairpins. For this check, we use restriction enzymes. Figure 18 in the methods chapter shows where the enzymes cut the DNA. If AseI or MluCI cut the structure, this suggests that the folded state forms. In the same way, BsRBI shows whether the extended state forms. Graph 31a shows the results of the experiments with the construct and enzyme MLuC1. This enzyme cuts A/A' in one of the loops if this part of the structure is double-stranded. Looking at graph 31a and comparing with the situation where no enzyme is added shows that three peaks emerge at new locations. One of the peaks is slightly further than the original peak and two peaks a lot further in the gel. Looking back at figure 18, this can be explained by looking at the stability of the construct after a cut. Although the stem of the hairpins stabilizes the hairpins, this original stability reduces after the cut, due to the lower number of base pairs. So after dissociation,

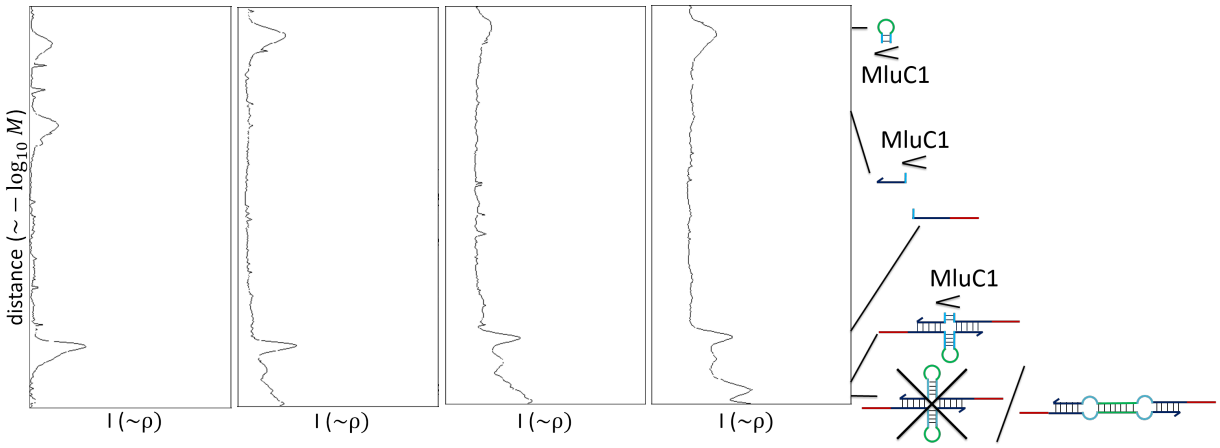


(a) *X*, $N_H = 25$. (b) *XY*, $N_H = 25$. (c) *X*, $N_H = 10$. (d) *XY*, $N_H = 10$. (e) *X*, $N_H = 7$. (f) *XY*, $N_H = 7$.

Figure 30: *Gel electrophoresis results of adding only strand X or both strand X and Y for the constructs with $N_H = 7$, $N_H = 10$ and $N_H = 25$. This gel suggests that for all constructs at 4°C, strand X and Y bind, forming a bigger structure that moves a shorter distance. The used gel is a 12% PAGE gel, and the used buffer is NEBuffer CutSmart.*

there are three pieces of DNA. The peak closest to the original location represents the big hairpin stem part, the second peak the smaller hairpin stem part and the piece that travels the biggest distance the small (linearized) loop of DNA.

Graph 31b, 31c and 31d show the result of adding the same enzyme to strand X and Y at varying temperatures. This time, two new peaks emerge. Now the construct does not break in three pieces since strand Y stabilizes the stem part of strand X. So one of the peaks corresponds to the small loop piece and the other peak to the rest of the construct. The relative intensities of the peaks only decrease slightly with increasing temperature, suggesting that at all of these temperatures hairpins are formed. At the bottom of the graphs, a peak

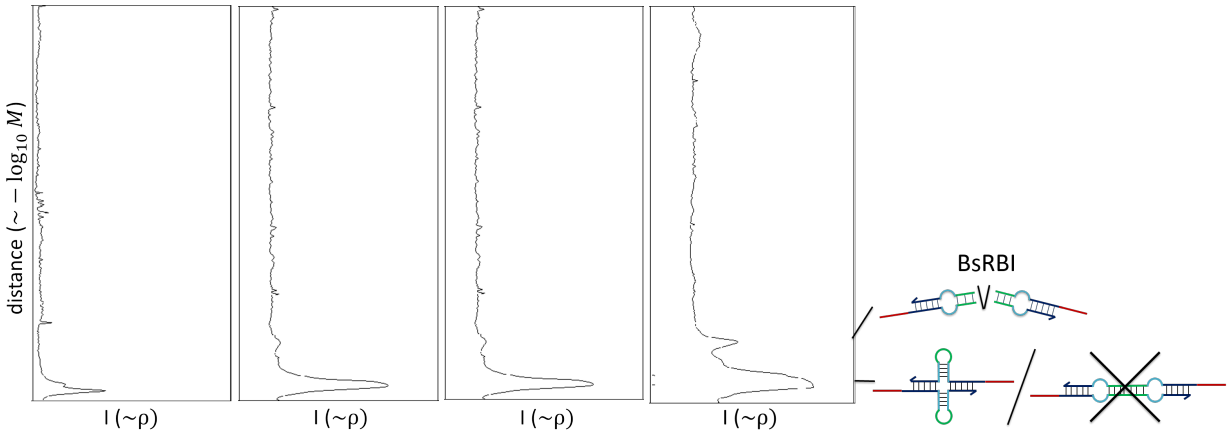


(a) *X*, $T = 25°\,C$ *Enzyme MLuC1.*

(b) *X and Y*, $T = 25°\,C$, *Enzyme MLuC1.*

(c) *X and Y*, $T = 37°\,C$, *Enzyme MLuC1.*

(d) *X and Y*, $T = 50°\,C$, *Enzyme MLuC1.*



(e) *X and Y*, $T = 25°$.

(f) *X and Y*, $T = 25°$, *Enzyme BsRBI.*

(g) *X and Y*, $T = 37°$, *Enzyme BsRBI.*

(h) *X and Y*, $T = 50°$, *Enzyme BsRBI.*

Figure 31: *Gel electrophoresis results of adding only strand X or both strand X and Y of construct* $N_H = 25$*, in combination with enzyme MLuC1 or BsRBI at three temperatures. The peaks correspond to the pieces of strand displayed in figure 18. The used gel is a 12% PAGE gel and the used buffer is NEBuffer CutSmart.*

can be observed at the hight corresponding to the full, uncut construct. So either part of the constructs is in a different state (for example the extended state) or that not all constructs where cut before the solution was added to the gel.

We can confirm the observations by comparing the lengths of these pieces of strands with the closest DNA ladder strand (see appendix for figures). The parts of the single DNA strand have lengths of 20, 31 and 91 bases. The pieces travel respectively approximately the same distance as the ladder strands corresponding to 30 bases, $40 - 50$ bases and 150 bases. The parts of full construct have lengths of 20 and 214 bases and travel the same length as the DNA ladder strands of 30 and 400 bases. The distances traveled by the smaller pieces is close to the expected distances, but slightly less far. A possible explanation is a difference in resistance of single-stranded and double-stranded DNA. The difference is more significant for bigger structures than for small structures. Possibly, the Holiday-junction causes higher friction.

Graph 31f shows the result of adding enzyme BsRBI, which cuts B/B'. At $T = 25°$ C, the gel is similar to the gel where the construct without the enzyme (Gel 31e). This observation suggests that the extended state is (almost) not formed at 25°C. So together with the other results, these results show that without an applied force, at 25°C, the folded state is formed while the extended state is (almost) not formed.

When a high enough force is applied, the extended state should form before the two sides of the DNA catch bond break. It is not possible to do gel electrophoresis experiments with a construct under force, but there are different ways to make the extended state more likely. The folded state is more stable than the extended state at low temperatures due to the lower enthalpic contribution to the free energy. However, entropically, the extended state is favored over the folded state. Free energy calculations using the Nearest-Neighbor model show that at temperatures of around 45°C the free energy is equal in both states. So from around this temperature, a significant portion of strands is extended.

Graph 31f, 31g 31h shows the effect of adding enzyme BsRBI to both strands at three different temperatures. At 25°C only a tiny extra peak is observable, so the extended state is only present in minimal amounts. Increasing the temperature to 37°C and 50°C results in a more significant peak. The distance traveled by the equally sized, 142 bases long pieces corresponds to the distance traveled by the ladder strands of 150 bases, as expected.

Figure 32 shows the relative intensities over temperature, at the positions in the gel corresponding to the sizes of parts of strands. To compare the intensities, the areas inside the peaks are normalized to the full sum of the intensities of all peaks and rescaled with the intensity at 25°C. This figure shows that the number of strands in the folded state decreases and the number of strands in the extended state increases with temperature. So not only the folded state but also the extended state can be formed.

We note some caveats with the data in this section. A first remark is that the relative intensity of different peaks does not necessarily represent the relative number of strands within a gel. The longer a DNA strand, the more of the used Sybr gold stain attaches. Also, there is a difference between the amount of stain attaching to single-stranded and double-stranded DNA. Strands with equal lengths can be compared. In some of the gels, we observe a lot of background fluorescence. After staining a gel, the stain not attached to DNA should be washed away. The background signal in these figures could be caused by small amounts of stain still present in the gel when the picture is taken.

At the location where DNA is initially added to the gel (see pictures in C), a high amount of fluorescence is measured. To make small parts of DNA such as the small piece in figure 18

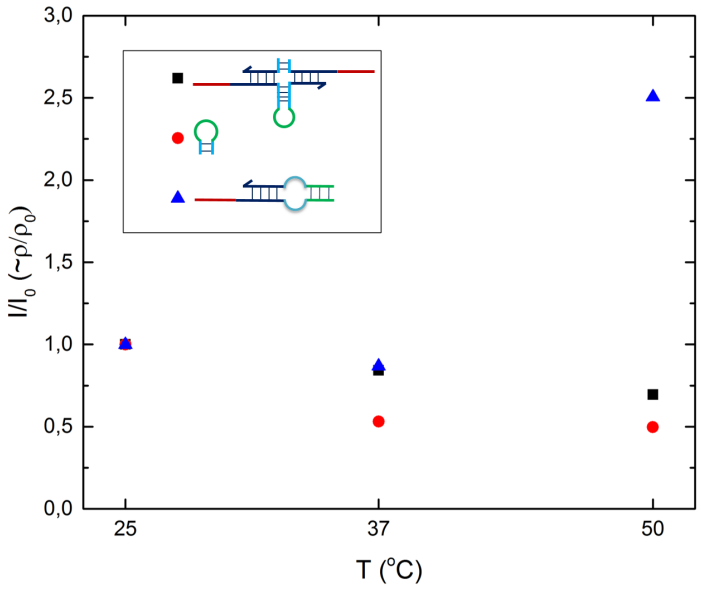


Figure 32: *Relative intensities, compared to the intensities at* $25°C$ *of the DNA strands resulting from DNA cuts. More measurements are needed for the uncertainties. This figure suggests the number of strands in the folded state decreases with temperature while the number of strands in the extended state increases with temperature.*

observable, we added a big amount of DNA in the gel. Some of the DNA gets stuck at the start of the gel. Possibly, large DNA aggregates form at the beginning of the gel. Another remark that should be made is that it is possible that a DNA strand switches from the folded to the extended state (or vice versa) during the time that the added enzymes are active. So if all of the DNA strands are cut, this does not mean all strands are in the state observed with gel electrolysis all the time.

To conclude, the gel electrophoresis experiments suggest that the X and Y strand of DNA catch bonds form a construct. At room temperature, the stem of the hairpins is cut by restriction enzymes while the B/B'-region is not cut, suggesting the constructs are in the folded state. Raising the temperature shows that also the B/B'-region, so also the extended state, can be formed.

4.1.2 Force induced state-switching

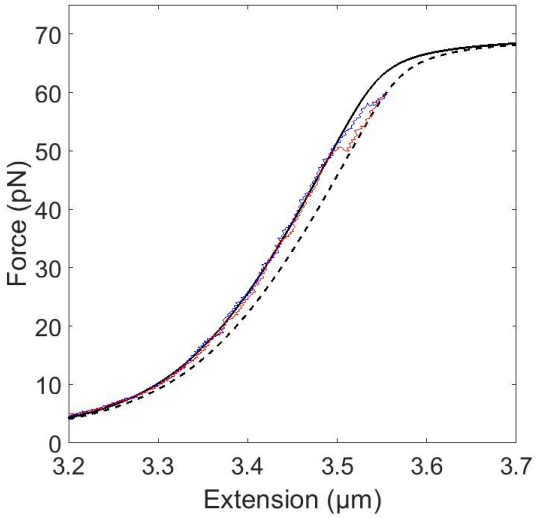
Now we know that both the folded ground state and extended state of the DNA catch bond form, we take a look at whether it is possible to switch from the folded to the extended state by applying a force. Optical tweezers are used to apply a force to the DNA catch bond construct. Here, the results of both the constant force and force-extension measurements are discussed.

Force-extension measurements We first look at the force-extension measurements. As the applied force increases, the end-to-end distance of the DNA strand increases. In the case of DNA catch bonds (including the two spacers of 5024 base pairs), there are two possible dependencies: One for the folded construct and one for the extended construct. In the extended state, the backbone length is longer than the backbone length in the folded state, with a difference equal to the length of the extended hairpins. This additional contour length changes the force-extension curve. So from the exact force-extension curve can be deduced whether a construct is in the folded state or the extended state.

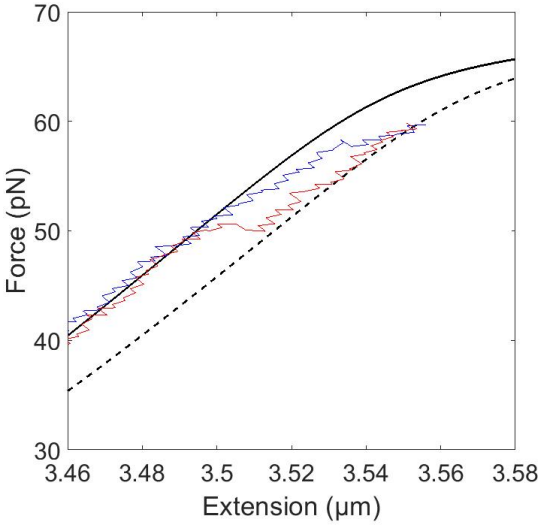
However, we cannot determine the exact force-extension curve. Section 3.3 already describes that the DNA tethers can attach anywhere on the used colloids. So for every distance between the centers of the colloids, there is a range of possible end-to-end lengths of the DNA tethers. The length difference caused by switching from the folded to the extended state is relatively small in comparison to this variation. So the absolute extension cannot be used to deduce whether a DNA catch bond is in the folded and the extended state. Also, the length difference is relatively small in comparison to the length of the spacers. As a result, the force-extension curves are similar for the folded and the extended state. So also the shape of the measured force-extension curve cannot be used to deduce whether a DNA catch bond is in the folded and the extended state. The state of the DNA catch bonds can be observed when a switch of states is observed. In that case, both force-extension curves are followed in part of the data.

Figure 33 shows an example of a set of data. This set of data consists of multiple consecutive pulls with the same tether. In each of the force-extension curves, we can see that the experimental data follows the theoretical force-extension curve for the folded construct for low forces. As the force increases, the curve switches from the curve corresponding to the force-extension curve of the folded state to the curve corresponding to the extended state. When the distance between the colloids reduces, the curve follows the force-extension curve of the extended state, until it switches back to the curve of the folded state. These transitions suggest that force-induced state switching is possible. The average forward switching force measured in these experiments is 53 pN. Our model, as described in chapter 2, predicts forward transitions at 51 pN, agreeing well with the experiments. The average backward switching force is 50 pN. These transitions are discussed further in section 4.1.3. More experiments must be performed to get accurate standard deviations and errors.

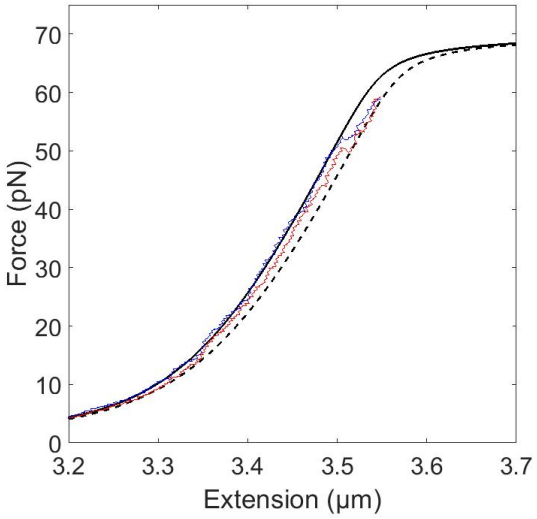
Not all measurements followed the theoretical force-extension curves. Figure 34 shows two examples of common other measurements. A possible explanation of these measurements is that multiple DNA catch bonds cluster together and stabilize each other. Decreasing the concentration of tethers did not eliminate these observations. Possibly, DNA catch bonds cluster together even at low concentrations, because the unbound parts of constructs are complementary to the unbound parts of other constructs. The outlook of the thesis gives some suggestions for cleaner experiments to confirm the observations in this section.



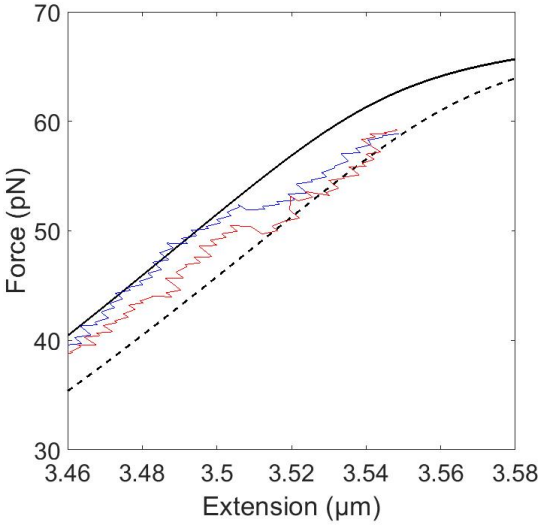
(a) *Force-extension curve of first experiment.*



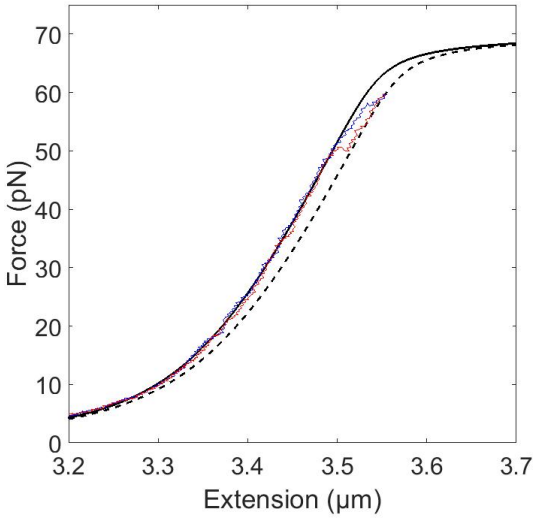
(b) *Zoom of state switch of third experiment.*



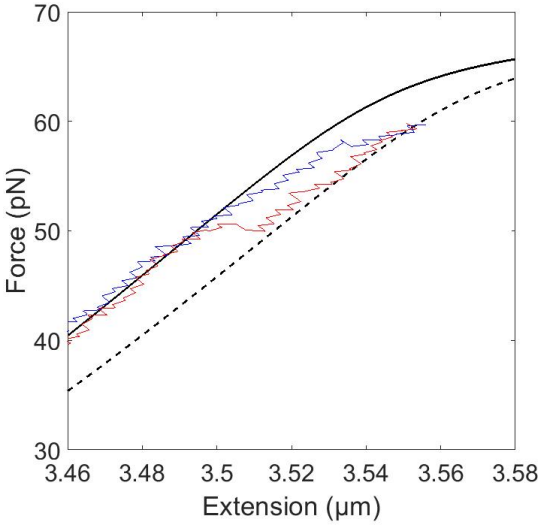
(c) *Force/extension curve of second experiment.*



(d) *Zoom of state switch of second experiment.*

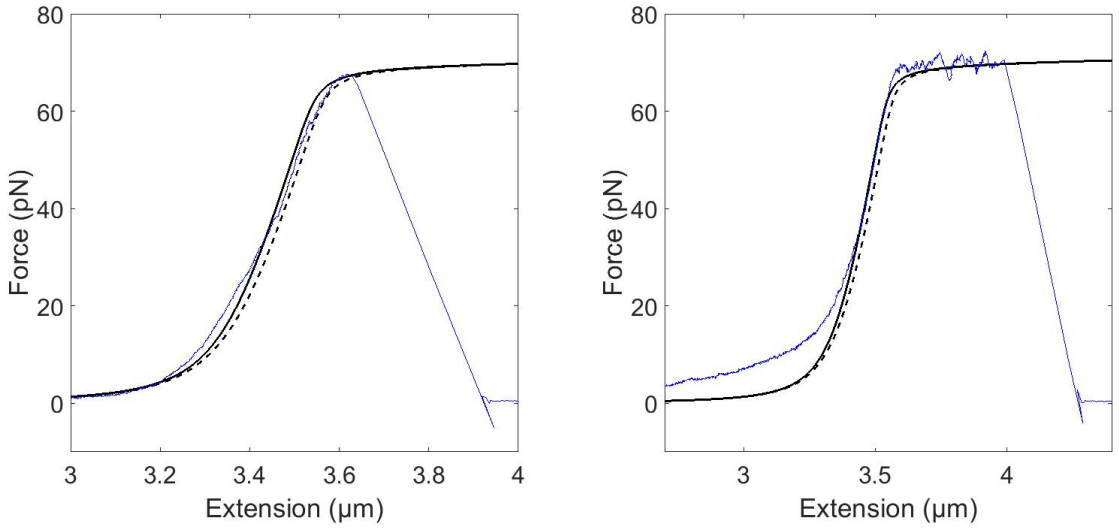


(e) *Force/extension curve of third experiment.*



(f) *Zoom of state switch of third experiment.*

Figure 33: *Three consecutive force-extension experiments, from the same set of colloids (sequence 1 and 2 in appendix A). Blue data correspond to increasing force and separation and the red data to relaxation. The colloids in these experiments are moved apart with a constant rate of* $0.1\mu m/s$.



(a) *Over-stretching and showing a potential switch.* (b) Over-stretching, but no switch observable.

Figure 34: *Two examples of force-extension curves of experiments where the measured force-extension curve does not align with the theoretical force-extension curves.*

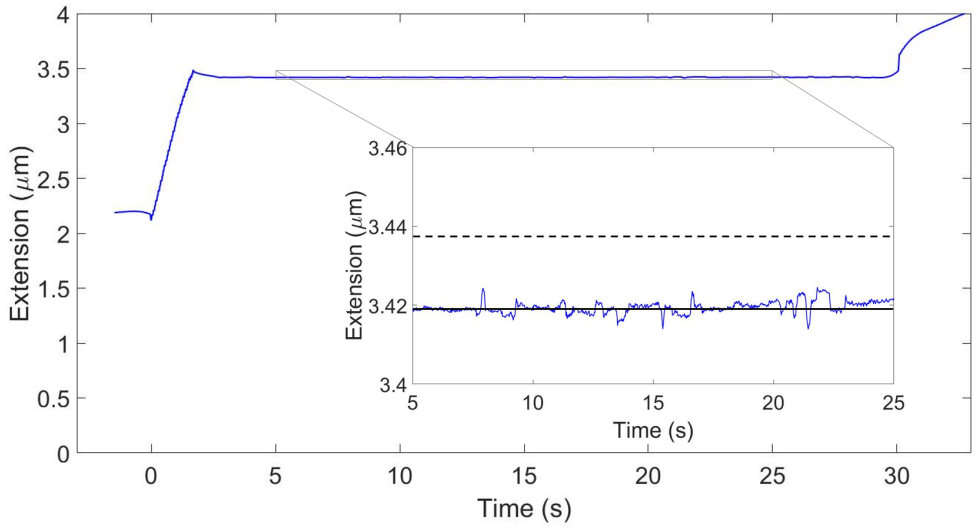
Constant force measurements More information can be gathered from the constant force measurements. One might argue that when the force is chosen close to a critical force, the construct should be in the folded state for part of the time and in the extended state for part of the time. In equilibrium, the relative occurrence of a state is related to the force and the difference in free energy and length of the compared states:

$$\frac{\langle t^f_{fold}\rangle}{\langle t^f_{ext}\rangle} = 10^{\frac{1}{2}\Delta N} \exp\left(-\frac{f\Delta x}{k_B T}\right). \tag{65}$$

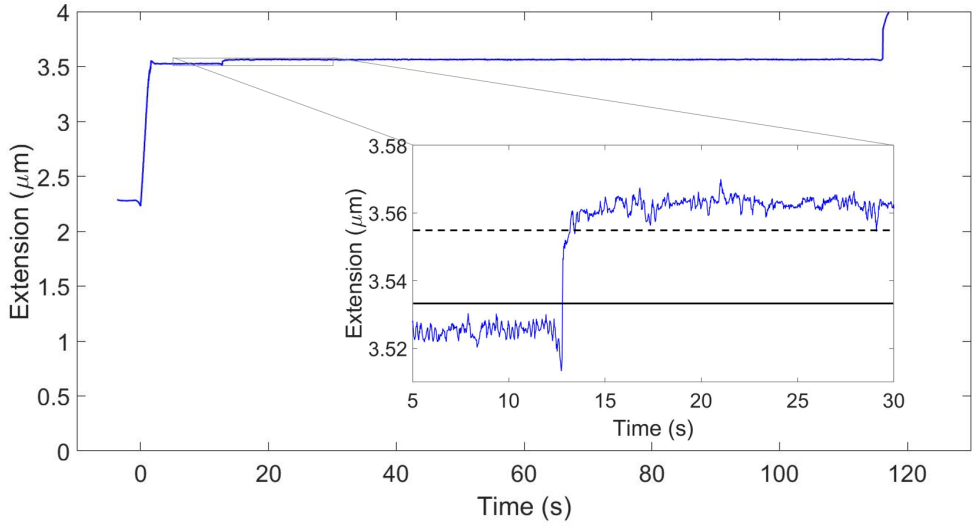
For equal times in the extended and folded state, the above equation gives a force of 32 pN (for sequence 3 and 4 in appendix A). So the reasoning above suggests that at this force, an equal amount of time is spent in the folded and the extended state. However, this reasoning is only correct if the total time spent in either of these states is long enough, which is not the case for the studied construct.

Every experiment starts at 0 pN, with the construct in the folded state. After the force is increased to the set constant force, two things can happen: The hairpin can open or the the handles of the construct can break, resulting in a dissociation of the catch bond. At low forces, including 32 pN, the average time for the opening of the hairpins is larger than the average time until the handles break. So the likelihood of observing a state switch is low. An estimation can be made of the relative times associated with switching, in the way as described in chapter 2:

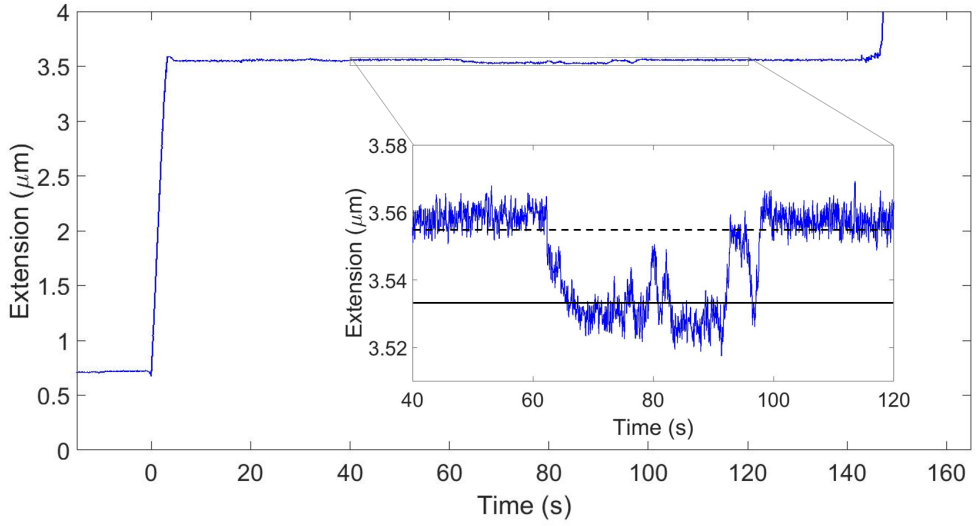
$$\frac{\langle t^f_A\rangle}{\langle t^f_H\rangle} = \frac{10^{\frac{1}{2}2N^{eff}_{N_A}-3} \exp\left(-\frac{f\Delta x_A}{k_B T}\right)}{10^{\frac{1}{2}N^{eff}_{N_H}-3} \exp\left(-\frac{\frac{1}{2}f\Delta x_H}{k_B T}\right)}. \tag{66}$$



(a) *No switching at 30 pN.*

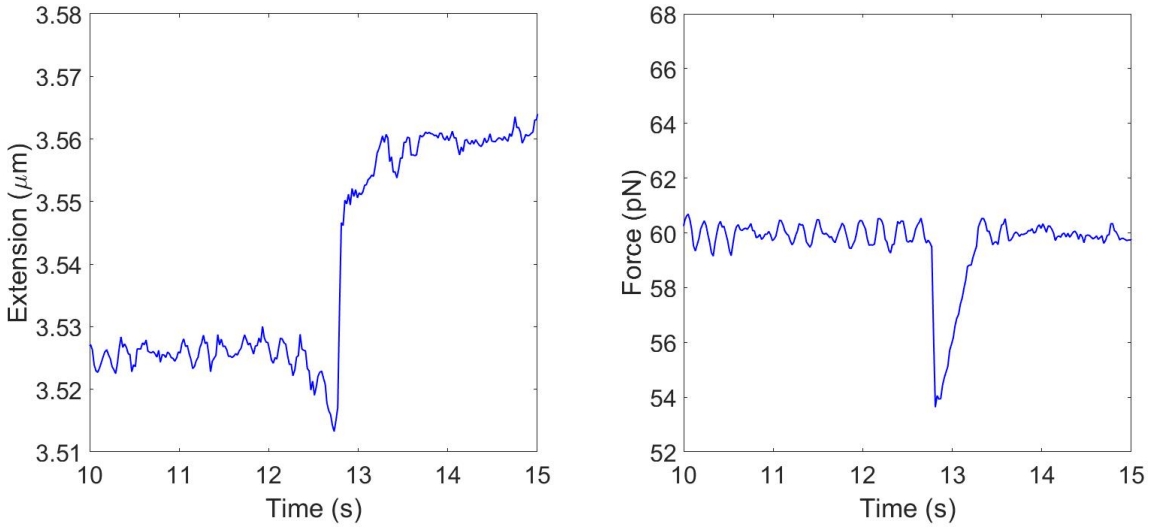


(b) *Forward switch at 60 pN.*



(c) *Forward and backward switches at 60 pN.*

Figure 35: *Examples of the extension over time of DNA catch bond construct stretched with a constant force of 30 pN and 60 pN. The zooms contain the theoretically predicted lengths in the folded (solid line) and the extended state (dashed line). Most measurements show no switching, some show forward switching and only the measurement shown in figure (c) shows backward switching.*



(a) *Zoom of extension/time.* (b) *Zoom of force/time.*

Figure 36: *Zoom of the extension-time and force-time of a transition from the folded to the extended state. Besides the jump in length in the extension-time graph, sometimes a valley is observable in the force-time graph.*

At 32 pN, the average time before the hairpins open is more than 700 times larger than the average time for the handles to break. So due to the instability of construct, which is needed for the DNA catch bond to function as a catch bond, state switching is unlikely to be observed at low forces. So it is very likely to observe only the folded state since the experiment started in the folded state.

At higher forces, the unzipping time becomes small. However, from a similar calculation as in equation 66 follows that for forces with small enough forward transition times, the backward transition times are high. So at high forces, the constructs will most likely stay extended after the initial switch.

Figure 35 shows some examples of extension-time curves of constant force experiments. During most of these experiments, I observed no state switching. In these measurements, either the construct stays in the folded state, or the construct already switches to the extended state during the acceleration. In a couple of high force experiments I observed a forward transition, and in a single measurement, I observed a backward transition. The difference in end-to-end length between the longer and shorter state observed in experiments is on average slightly larger than the theoretically predicted difference in length. This difference can be partly, but not entirely, explained by the B/B'-region staying single-stranded. Section 4.1.3 gives more information about the formation of the stabilizing duplex.

Another indication for state switching can be observed in the force-time curves. During a switch, the end-to-end length of the DNA strand changes. As a result, the force on the construct changes. In some cases, I observed a valley or hill in the force-time plot. The opposite happens during a switch from the extended to the folded state. The feedback mechanism for the force quickly corrects these changes in applied force. The predicted increase and decrease in force can be calculated used the chain model. The equations are kept equal, except the

end-to-end length is increased for extension transitions and decreased for a folding transition. This calculation results in an expected force decrease of 5.2 pN for an extending transition and a force increase of 3.6 pN for a folding transition (from 60 pN). The observed decreases and increases in force are generally smaller than these forces. Figure 36 shows an example of a valley in the force-time plot and the corresponding extension-time plot.

To summarize, state-switching is observed both in constant force experiments and force-extension experiments, but not in all experiments. The different observations in other simulations could be caused by clustering of DNA catch bonds. More experiments are needed to explain these observations and to determine the critical force more accurately.

4.1.3 Hybridization of stabilizing duplex

The optical tweezers results of the previous section suggest the hairpins of the DNA catch bond open. For the functionality of the DNA catch bonds, the extra, stabilizing duplex should form after the hairpins open. In this section look back at the optical tweezer results to see whether the stabilizing helix formation is likely happening.

The average time for a transition between each of the possible states of DNA catch bonds can be calculated using free energies, end-to-end lengths and forces corresponding to all intermediate states. The higher the free energy at the top of the barrier between two states, the higher the average transition time. For the transition from the extended state to the folded state, the B/B'-region should dissociate. This is the main barrier for backward transitions when no force is applied:

$$\langle t_B^0 \rangle = 10^{\frac{1}{2}N_{N_B}^{eff}-3} \text{ [s]}. \tag{67}$$

The time needed to cross this barrier decreases exponentially with force.

When you apply a force to an extended catch bond, there is an additional barrier due to the needed reduction in end-to-end length. The time needed to overcome this barrier is described by

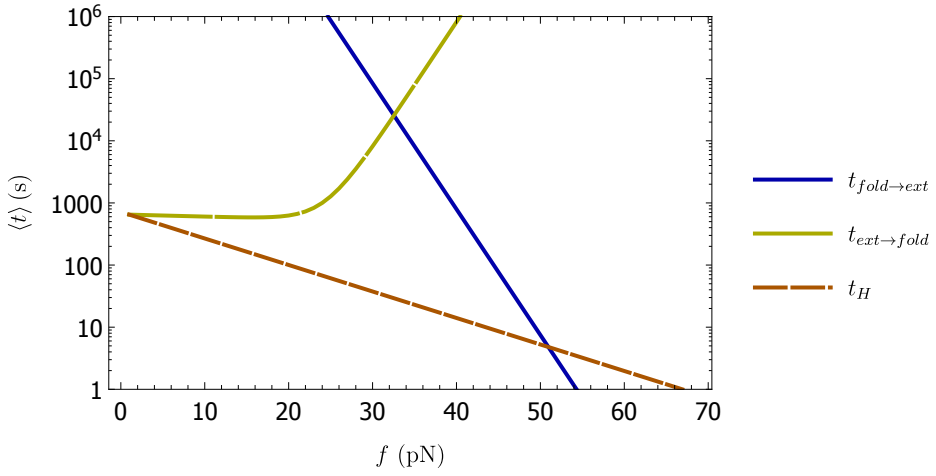
$$\langle t_x^f \rangle = 10^{-3} \exp\left(-\frac{f^{eff}\Delta x_x}{k_B T}\right) \text{ [s]}. \tag{68}$$

If the B-region does not hybridize, this is the only barrier between the extended and the folded state. Figure 37 shows the theoretical average transition times if the B-region hybridizes and if the B-region does not hybridize in the extended state.

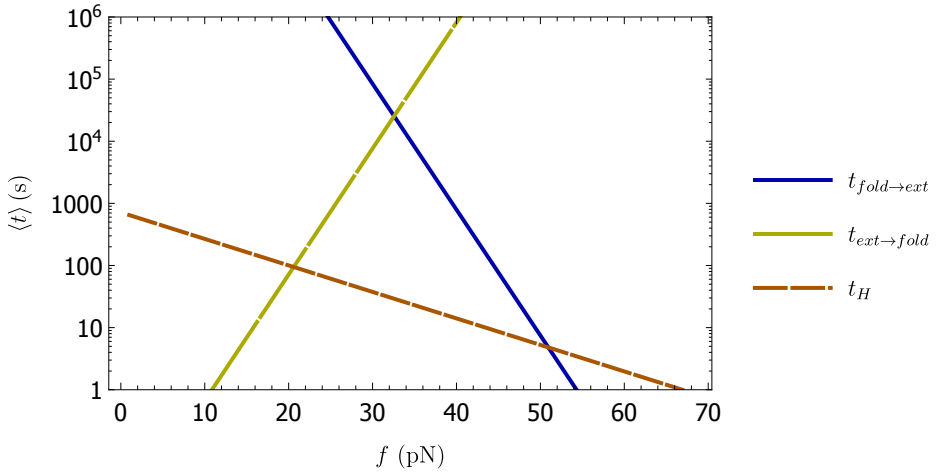
In the previous section, we saw the force-extension curves of optical tweezer experiments. During the reduction of the end-to-end length, the force-extension curve switched from following the curve corresponding to the extended state to the curve corresponding to the folded state. This suggests the construct switches back at a finite force, which is unlikely to happen when the B/B' region hybridizes. So this data suggests the B-region does not hybridize during these experiments.

Possibly, the B-region does not hybridize due to the limited time spent in the extended state. A possible way to check this in future experiments is by running a force-extension experiment where the force increases until the force-extension curve switches from following the curve corresponding to the folded state to the extended state. Next, the force is kept constant for some time before the force reduces. The B-region formed during the time in the extended state if there is no backward switch observable (until low forces). The intended catch bond effect is absent if the B/B'-region still does not hybridize within the rupture time of the construct. A possible solution to this problem would be to increase the length of the B-region to increase the likelihood of hybridization.

So to conclude, the optical tweezer experiments suggest that the B-region does not hybridize. More experiments are needed to check whether the B-region can form under other circumstances. In future experiments, the formation of the extra region can be tested with longer times at high forces and with longer B-regions. These adjustments increase the likelihood of hybridization of the B-region.



(a) *B binds to B′.*



(b) *B does not bind to B′.*

Figure 37: *Transition times from the folded to the extended state and from the extended to the folded state and the lifetime of the handles.*

4.1.4 Increasing lifetime

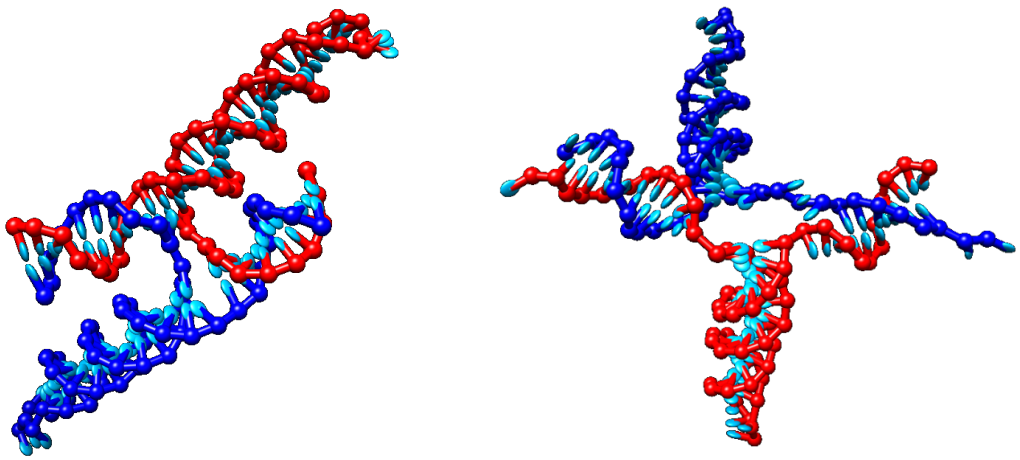
In the previous section, we saw that more experiments are required to confirm whether the stabilizing region forms. Possibly, the constructs must be altered to fit this requirement. In this section, we want to find out whether the lifetime of the DNA catch bonds increases, assuming the stabilizing region forms. Using simulations, we can enforce the state of the constructs to be in the folded or the extended state, making it possible to compare these states, even if it is unclear whether the stabilizing duplex forms.

Before we can start using simulational results, we should check that results are as expected. Figure 38 shows snapshots of the constructs in the folded state. In the initial, relaxed state when no force is applied, the hairpins and handles align in two parallel arms. When a force is applied, the hairpins are quickly tilted to maximize the distance between the pulled apart ends. So visually, the behavior agrees well with our expectations.

I test the simulation method further by running simulations and comparing with literature. An example of such a test is the dissociation time of short pieces of double-stranded DNA while being shear unzipped with a force of 50 pN. Figure 39 shows the average dissociation time as a function of the number of base pairs. As expected the dissociation time increases exponentially with the number of base pairs. I fitted the dependency according to:

$$\langle t^f \rangle = t_0 \, 10^{\alpha N_N^{eff}} \exp\left(\frac{f\Delta x}{k_B T}\right). \tag{69}$$

with:



(a) *Initial configuration.* (b) *Configuration during stretching.*

Figure 38: *Two examples of configurations during an simulation of the folded state. The red and blue color each represent one of the DNA strands.*

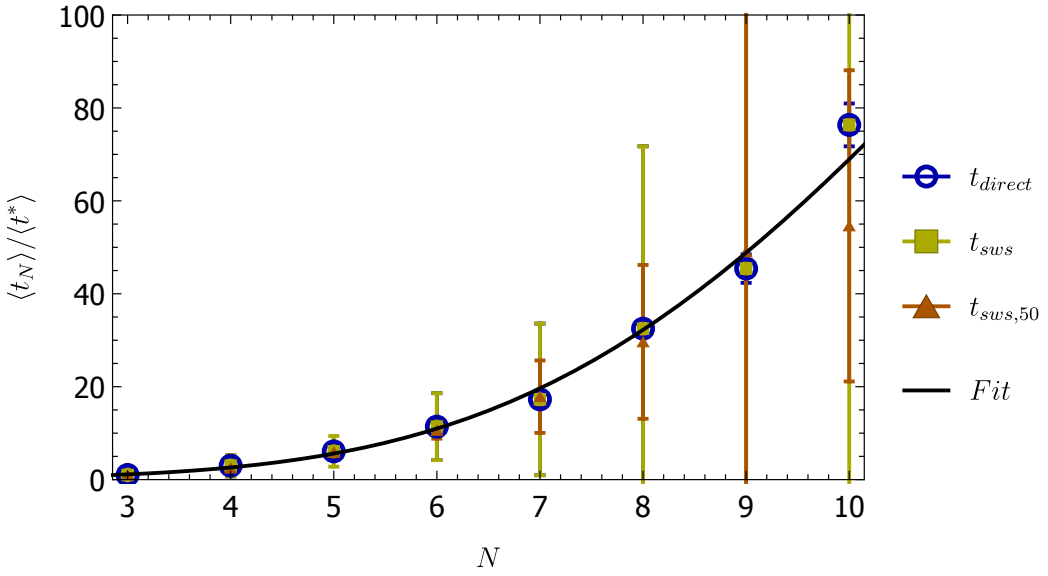


Figure 39: *Graph of average dissociation times of simple double-stranded DNA strand as a function of the length of these strands. The figure shows the results of direct measurements, Stairway Sampling measurements without pruned tracks and Stairway Sampling measurements with pruned tracks.*

$$\begin{aligned} N_N^{eff} &= 2\left(1+\chi^{-1}\tanh\frac{\chi}{2}N\right); \\ \Delta x &= \left(A + B\,N_N^{eff}\right)\ [\mathrm{nm}]\,. \end{aligned} \tag{70}$$

α, χ, A and B are the fitting parameters. The best fit gives $\alpha = 0.652 \pm 0.002$, $\chi^{-1} = 4.2 \pm 0.7$ base pairs, $A = 0.62 \pm 0.07$ nm and $B = 0.04 \pm 0.01$ nm. T. Strunz et al. found $\alpha = 0.5 \pm 0.1$, $A = 0.7 \pm 0.3$ and $B = 0.07 \pm 0.03$ using experimental data [28]. The fit suggests that the stability of the structure increases slightly more with the number of base pairs in the simulations than in the experiments. A and B agree with the experimentally found values. Hatch et al. [29] found $\chi^{-1} = 6.8$ base pairs. So the fit through the simulational data suggests that in the simulations, the force distributes over fewer base pairs than in the experiments.

Due to the wide variety, getting an accurate value for the average dissociation time takes many simulations. To reduce the computational time needed, I used stairway sampling as described in section 3.7. Figure 39 shows results using Stairway Sampling agree with simulational results of direct simulations. For the same amount of measurements, the uncertainty is higher using stairway sampling than the results from the direct simulations, but the needed time per simulation is also much lower.

I tested the effect of the extra region in the extended state by comparing the average dissociation time in simulations of constructs constrained in the folded state and the extended state. However, the results are not reliable enough to be presented here due to differences between simulational results and experimental results. For the data in figure 39, we saw that the constants found by fitting theoretical curves through our simulation results gives different values for the adjustment length than fits through experimental data. So in experiments, the force is distributed over fewer base pairs than in the experiments. The distribution of forces

is vital for the catch bond effect, so such a disagreement makes simulations less informative. Section 4.2.3 gives an example of how the simulational results of the full DNA catch bond disagree with the experiments.

Future simulations could be improved by changing the potential used for the backbone of DNA. By increasing the stiffness of these potentials, the force propagates further through the backbone. The backbone stiffness should be tuned until the adjustment length agrees with the lengths found in experiments. Of course, also the other characteristics of DNA should be checked after this change.

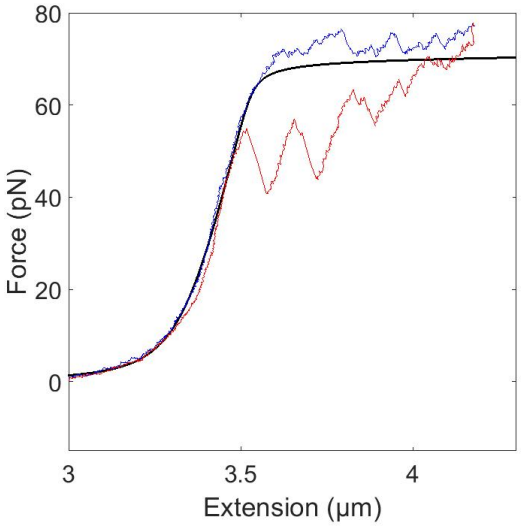
4.2 Lifetime-force dependency

During the research reported in this thesis, I used three different methods to determine the lifetime-force dependency of DNA catch bonds. The average lifetime must be measured at multiple forces to measure the catch bond effect. At each of these forces, a high number of measurements is required to get an accurate prediction of the lifetime. I used two types of experiments in an attempt to achieve this goal: Constant force optical tweezer experiments and constant force centrifugal microscopy experiments. In this section, I discuss why these experiments did not provide the needed data to conclude whether or not the DNA construct has a range of forces where the bond lifetime increases with force. Finally, this section discusses the results from the simulations.

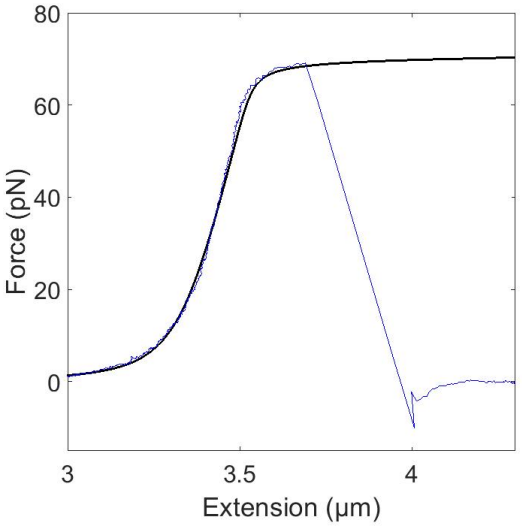
4.2.1 Optical Tweezers

As discussed earlier, optical tweezers can be used to apply a constant force to DNA catch bonds. The lifetime of the DNA catch bonds can only be measured accurately if the other connections used in such an experiment are strong enough. To check this, we first take a look at the force-extension curves of DNA constructs without the DNA catch bonds. The used tethers consist of the two $2 \cdot 5024$ base pair spacers, which will later be used in combination with the DNA catch bonds. Figure 40 shows two examples of resulting force-extension curves.

The second experiment shows an example of an experiment where the tether breaks. After the break, the used feedback mechanism reduces the force. The colloid in the optical trap oscillates for a short time around it is equilibrium position. At the same time, the distance



(a) *First experiment with increasing (blue) and decreasing (red) force.*



(b) *Second experiment where the tether breaks during the experiment.*

Figure 40: *Two examples of force-extension curves of the DNA spacers ($2 \cdot 5004$ base pairs) without the DNA catch bond construct. The second displayed experiment was performed with the same tether as the first experiment. The colloids in this experiment are moved apart with a constant rate of $0.1 \mu m/s$*

between the particles increases without a counteracting force. This experiment was performed with the same colloids and tether as the first experiment. In most cases, it was possible to extend the tethers multiple times to well within the over-stretching regime before the tethers break. So the connections between the colloids and the DNA are strong enough to withstand the force needed to overstretch DNA. This means that the DNA catch bonds can be stretched up to the overstretching force with a low chance that one of the other connections breaks.

Now we know that the other connections are stable enough, the lifetimes of DNA catch bonds can be measured. However, the problem with this method is that it takes a long time to gather enough data to find the average dissociation times. After a successful measurement, it takes a long time for one side of the construct to find its counterpart and fully rebind. So after every experiment, the process of preparing an experiment must be repeated. This process includes capturing a NeutrAvidin coated colloid, washing the channel, trapping a DNA coated colloid and rewashing the channel. Finally, the optical trap must be re-calibrated. A low concentration of DNA tethers is used, to suppress the chance of interaction between tethers. A result is that there are colloids without any attached DNA tethers. So sometimes this process must be repeated without any successful measurements. So performing many experiments takes much time.

Even if one would spend enough time to gather this data, it is not certain that a catch bond effect would be observed. As discussed in section 4.1.2, a relatively large portion of the force-extension measurements show unexpected behavior. In constant force measurements, it is not possible to see the force-extension curve, so it is impossible to check whether a measured construct behaves as expected or not. The constructs behaving differently than expected could blur the catch bond effect to such an extent that a catch bond effect would not be observable anymore in the lifetime-force dependency. So before a large number of constant force experiments is performed, the cause of the experimental results should be found and eliminated. The outlook of this thesis suggests how to improve the construct and experiments.As discussed earlier, optical tweezers can be used to apply a constant force to DNA catch bonds. The lifetime of the DNA catch bonds can only be measured accurately if the other connections used in such an experiment are strong enough. To check this, we first take a look at the force-extension curves of DNA constructs without the DNA catch bonds. The used tethers consist of the two $2 \cdot 5024$ base pair spacers, which will later be used in combination with the DNA catch bonds. Figure 40 shows two examples of resulting force-extension curves.

The second experiment shows an example of an experiment where the tether breaks. After the break, the used feedback mechanism reduces the force. The colloid in the optical trap oscillates for a short time around it is equilibrium position. At the same time, the distance between the particles increases without a counteracting force. This experiment was performed with the same colloids and tether as the first experiment. In most cases, it was possible to extend the tethers multiple times to well within the over-stretching regime before the tethers break. So the connections between the colloids and the DNA are strong enough to withstand the force needed to overstretch DNA. This means that the DNA catch bonds can be stretched up to the over-stretching force with a low chance that one of the other connections breaks.

Now we know that the other connections are stable enough, the lifetimes of DNA catch bonds can be measured. However, the problem with this method is that it takes a long time to gather enough data to find average dissociation times. After a successful measurement, it takes a long time for one side of the construct to find its counterpart and fully rebind. This means that after every experiment, the process of preparing an experiment must be

repeated. This process includes capturing a NeutrAvidin coated colloid, washing the channel, trapping a DNA coated colloid and washing the channel again. Finally, the optical trap must be re-calibrated. A very low concentration of DNA tethers is used, to suppress the chance of interaction between tethers. This also means that there are colloids without any attached DNA tethers. So sometimes this process must be repeated without any successful measurements. So performing many experiments takes a lot of time.

Also, even if one would spent enough time to gather this data, it is not certain that a catch bond effect would be observed. As discussed in section 4.1.2, a relatively large portion of the force-extension measurements show unexpected behavior. In constant force measurements, it is not possible to see the force-extension curve, so it is impossible to check whether a measured construct behaves as expected or not. The constructs behaving differently than expected could blur the catch bond effect to such an extend that a catch bond effect would not be observable anymore in the lifetime-force dependency. So before a large number of constant force experiments is performed, the cause of the experimental results should be found and eliminated. A suggestion is provided in the outlook of this thesis.

4.2.2 Centrifugal Microscopy

Centrifugal microscopy provides another way to apply a constant force to DNA constructs. In these experiments, the centrifuge accelerates to the desired rotational velocity, after which the rotational velocity is kept constant. The advantage of this method is that measurements can be performed in parallel. Also, by using Nanoswitches, experiments can be repeated quickly. So this setup solves the problem of the needed time to gather data with optical tweezers. We used this method to apply constant forces to DNA catch bond constructs. Histogram 41

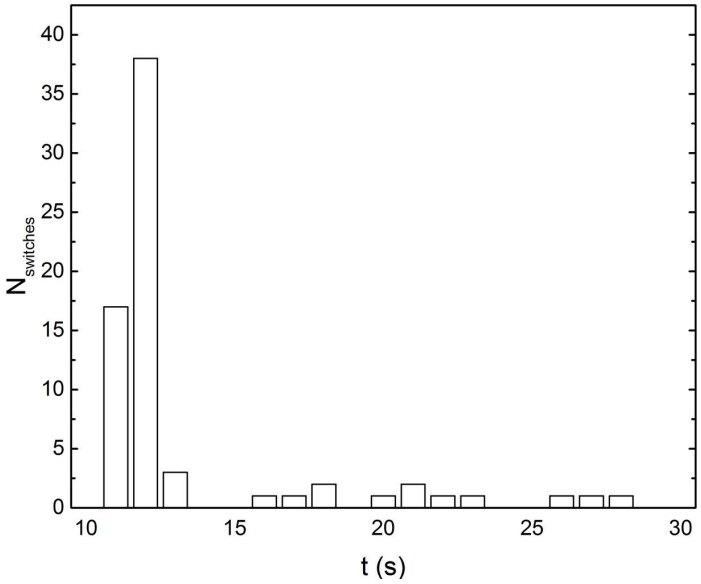
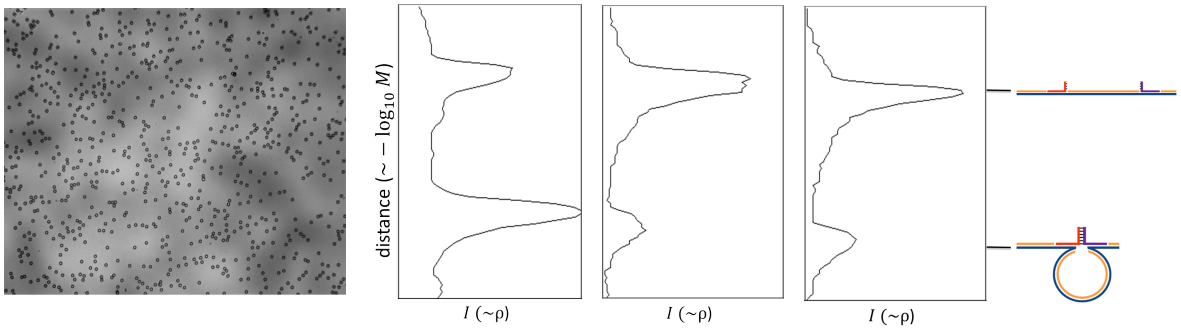


Figure 41: *Combination of data from three experiments at $f = 29 \pm 2$ pN, including the time where the centrifuge is still accelerating. The figure shows that most of the unbinding happens before the constant force is reached (15s). This limits the possibility to use this method to find the lifetimes at specific forces.*



(a) *Example of a frame from an experiment with DNA catch bonds.*

(b) *29 base pair Zipper construct.*

(c) $N_H = 10$.

(d) $N_H = 25$.

Figure 42: *To suppress multi-tethering, the surface density of tethers is lowered in the DNA catch bonds experiments. (a) A frame of an experiment at low concentration. Figure b, c and d show the results of Gel electrophoresis experiments with Nanoswitches with the DNA catch bonds and the zipper construct. Due to the higher friction with the gel for folded Nanoswitches, a closed Nanoswitch moves less far than an open Nanoswitch. The top peak represents the open Nanoswitches and the bottom peak the closed Nanoswitches.*

shows an example of the dissociation time distribution found in three centrifuge spins. Most of the dissociation events happen 10 to 15 seconds after the start of the experiment. For higher times, a relatively low number of dissociations are observed.

Unfortunately, the data from these measurements cannot be used reliably to find the lifetime-force dependency of the DNA constructs. The reason is the too low force rate of the setup. The acceleration of the centrifuge is low in the first couple of seconds and reaches its maximum at the end of the acceleration period with an acceleration of about 6 pN/s. As a result, the centrifuge is accelerating in the first 15 seconds (for the experiments shown in figure 41). So most of the breaking events happen before the maximum force is reached. To get an accurate measure of the lifetime-force dependency, the acceleration time should be so small that it can be neglected, which is not the case in these experiments. These transitions before the constant force is reached do not only decrease the amount of statistics but also change the statistics. The state of the DNA catch bonds could already have changed before the constant force regime starts. Also, the effect of the acceleration is more significant for experiments at higher forces than for experiments at lower forces.

Higher force rates could be achieved in future experiments by increasing the mass of the colloids or improving the acceleration of the centrifuge. Nonetheless, the force rate can never be as high as in other setups such as optical tweezers. This limitation renders this method less useful to test the constant force response of DNA catch bonds. The method can, however, be used to measure the average breaking force of the DNA catch bonds in force ramp experiments. Appendix D presents some data of these measurements. For regular slip-bonds, the dissociation time increases with force rate. The limited data gathered for the DNA catch bonds indicates that the average dissociation time might be decreasing with the force rate. More force rate experiments must be performed before any conclusions can be drawn.

If we take a closer look at the data from the experiments, there are a couple of things to note. The fact that many of the catch bonds already dissociate during the acceleration

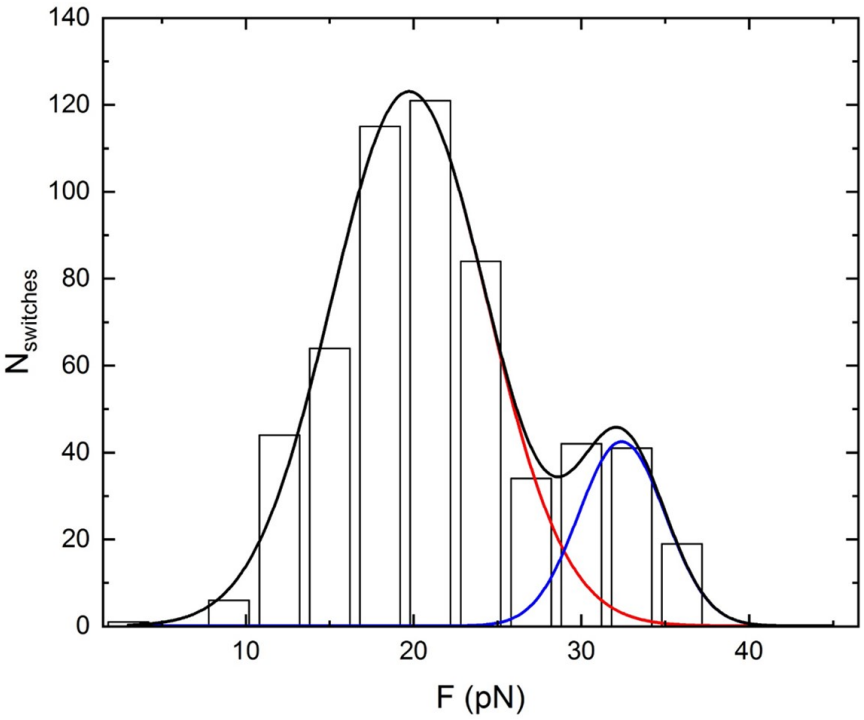


Figure 43: *Histogram of breaking forces from 1-2 pN/s force ramp experiments on DNA zipper strands of 29 base pairs. The data in this figure is a combination of five experiments at* $20°C$*. The average force and standard deviation are given by* 19.7 ± 0.2 *pN with a standard deviation of* 9.3 ± 0.6 *pN for first (Red) fitted Gaussian curve and* 32.4 ± 0.5 *with a standard deviation of* 5.2 ± 1.1 *pN for the second (Green) fitted curve.*

period is surprising if we compare the dissociation forces with the dissociation forces found in the optical tweezer experiments. Many of the measured dissociation times correspond to low dissociation forces in comparison to the dissociation forces found using optical tweezers, suggesting that the DNA catch bonds destabilize in the Nanoswitch setup. There are two ways to explain this destabilization. First, the entropy gain achieved by the opening of a Nanoswitch could contribute to the instability of the construct. Secondly, the structures could destabilize as a result of the acceleration. During the acceleration, the rotating motion of the centrifuge passes a resonance frequency. At this frequency, the sample vibrates slightly. These vibrations could lead to temporary higher forces applied to the DNA constructs, resulting in a destabilization.

Furthermore, only part of the constructs dissociates during the experiment. A possible cause for this low number of opening Nanoswitches is the attractive force between different catch bond structures. As described earlier in section 4.1.2, the B and B' regions of different DNA catch bonds are complementary, resulting in an attractive force between the different DNA catch bonds. As a result, the tethers could move together inside the sample and attach to the functionalized surface at the same location, increasing the chance on multi-tethering. To suppress this effect, I decreased the concentration of DNA. Figure 42a shows an example of a microscope view. The reduced concentration did not lead to experiments with more switches.

Another possibility is that not all Nanoswitches are closed at the start of the experiment. Unlike with the Optical Tweezer experiments, there is no way to check that bonds form. The Nanoswitches keep the colloids in the field of view, even if there is no connection between the two sides of the catch bond. So some of the catch bonds might already be opened before the experiment is started. To check the stability of the Nanoswitches with the DNA catch bonds, I performed gel electrophoresis experiments. Figure 42 shows the results of a gel electrophoresis run with Nanoswitches in combination with a simple zipper construct and the catch bond construct. In the folded state the gel hinders the movement of the Nanoswitches more than when they are open, so open Nanoswitches move further than closed Nanoswitches. The results suggest that a lower fraction of Nanoswitches is closed for the DNA catch bonds than for the simple zipper construct.

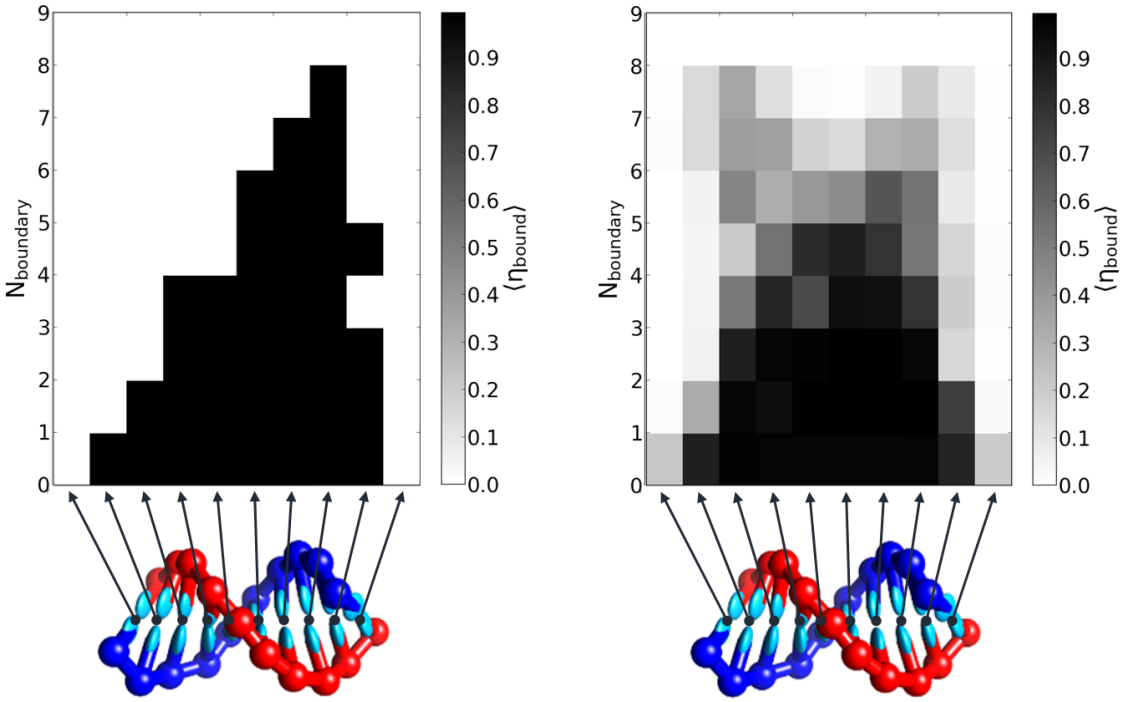
This observation is tested further by running a centrifugal microscopy experiment for the simple zipper construct and comparing the results with the results for the experiments with the DNA catch bond construct. Figure 43 shows the results of a force rate extension experiment with this construct. Fitting two Gaussian distributions gives an average force of 19.75 ± 0.24 pN. The second, smaller peak at a higher force is likely caused by multi-tethering. The measured force is higher than the average force of 14.6 ± 1.1 measured in previous experiments by D. Yang et al. with the same DNA sequence in the same setup (but in a different, tris(hydroxymethyl)aminomethane buffer with Magnesium ions) [40]. In these experiments, a much higher percentage of Nanoswitches open, giving additional support to the idea that the Nanoswitches with DNA catch bonds were open at the start of the centrifugal microscopy experiments.

4.2.3 Simulations

Simulations are a third way to find the average lifetime-force dependency. An advantage of simulations is that many lifetimes can be measured without consuming much time of the researcher (after programming). Also, the applied force can be set to a finite value instantly, eliminating problems with the acceleration time. Another advantage is that the locations of base pairs can be saved over time, giving information about the dissociation pathways.

Figure 44a shows a possible pathway from bound to unbound for a ten base pair long double-stranded DNA helix. The structure starts fully bound, and at each level, the number of base pairs reduces by one. Likely, the number of base pairs also increases during the simulations. Increases cannot be observed in this type of graphs, because the next configuration saved is the configuration with one base pair less than the previously saved configuration. So it is possible that the number of base pairs between step 5 and 6 first increases to eight before reducing to three. Figure 44b is the combination of one hundred thousand of these simulations. In this way, we can visualize the most likely pathways during the unzipping transition.

I applied the same method to the unbinding pathway of DNA catch bonds. For DNA catch bonds we expect two variations: One where the handles dissociate before anything else fully dissociates and the second where first the hairpins unzip (and the B-region hybridizes), before the connections between the two strands breaks. However, in all the simulations I performed, only the pathway where only the handles dissociate is observed. Increasing the force, reducing the length of the hairpins or reducing the length of the handles did not change the pathway. Figure 45 shows an example of an average pathway. If only one of the pathways is followed, there is no catch bond effect. So according to this data, there would be no range



(a) *One pathway.* (b) *Average pathway.*

Figure 44: *DNA can unzip in many pathways. Whenever the number of base pairs decreases with one base pair, the base pair locations are saved. (a) An example of a pathway from fully bound (bottom) to fully unbound. (b) Average pathway constructed from a combination of one hundred thousand pathways.*

of forces with a reversed lifetime/force dependency.

From the optical tweezer experiments, we already know that in reality, the catch bonds can switch between the folded and the extended state. So the simulations and experiments do not agree. This disagreement is not caused by a program error that prohibits unzipping; simulations of simple DNA helices show DNA can be unzipped and shear unzipped. Also, it is unlikely that Stairway Sampling changes the occurrence of the pathways. Unlike Umbrella Sampling, the energy landscape is not altered in Stairway Sampling.
Earlier we saw that the lifetime of double stranded DNA was accurately described by:

$$\langle t^f \rangle = t_0 \, 10^{\alpha N_N^{eff}} \exp\left(\frac{f \Delta x}{k_B T}\right). \tag{71}$$

with:

$$\begin{aligned} N_N^{eff} &= 2\left(1 + \chi^{-1} \tanh \frac{\chi}{2} N\right); \\ \Delta x &= \left(A + B\, N_N^{eff}\right) \text{ [nm]}. \end{aligned} \tag{72}$$

The best fit for the data was found using $\chi^{-1} = 4.2 \pm 0.7$ base pairs. In contrast, Hatch

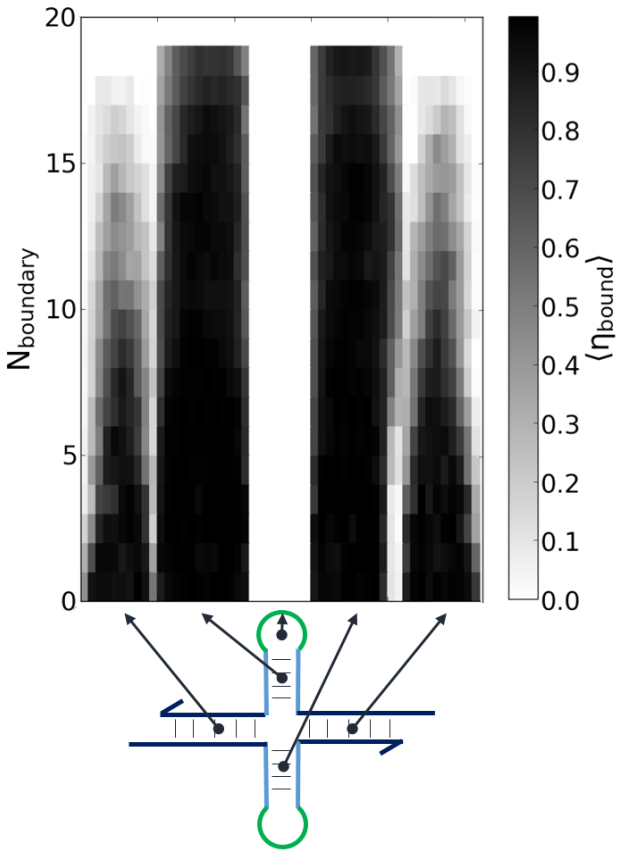
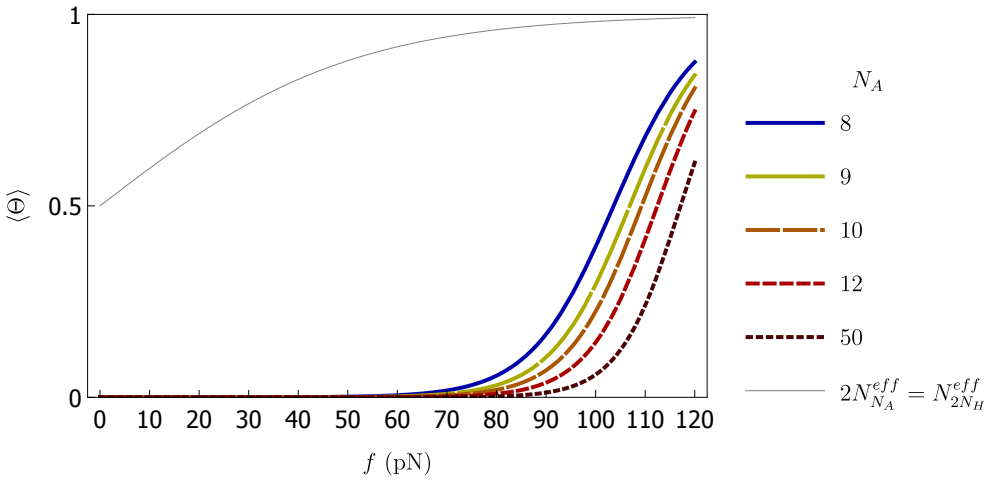


Figure 45: *Unbinding pathway of a DNA catch bond. In none of the performed simulations, the hairpins unzip.*

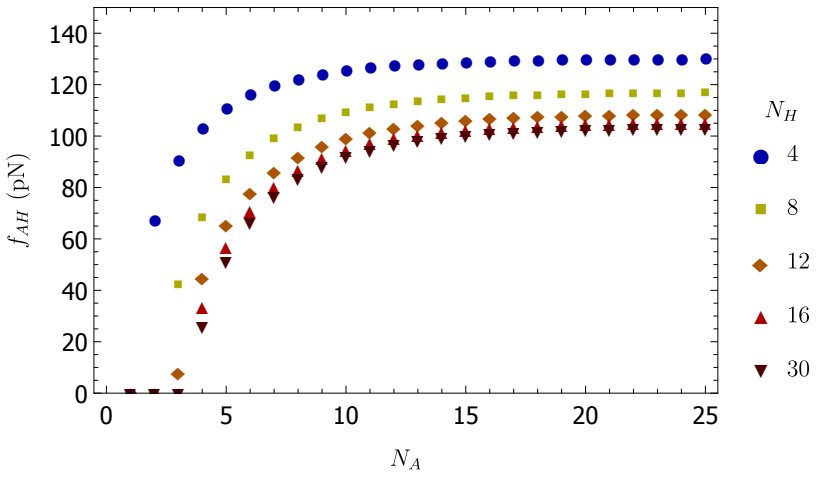
et al. [29] found $\chi^{-1} = 6.8$ base pairs gives the best fit for their experiments. So this data suggests that the force is distributed over fewer base pairs in the simulations than in the experiments. This disagreement could explain why no switching is observed in the simulations, while switching is observed in the experiments.

To test this statement, I recalculated the critical force using the parameters found in simple simulations. Figure 46 shows the results. The critical force increased to well within the overstretching regime for all the shown possible constructs. So this explains why no state-switching is observed around the state-switching force observed in experiments. Nonetheless, I also performed simulations at higher forces. Also in these simulations, no state-switching is observed. This observation can be explained by the timescale involved in the propagation of forces. When a very high force is applied to the ends of the DNA strands, the timescale associated with breakage of the hydrogen bonds becomes comparable to the timescale associated with the propagation of forces. So the handles already break before the hairpin is loaded.

So the simulations so far cannot be used to find the average lifetime-force dependency. In the future, the backbone potential could be improved, so simulations and experiments agree. Then, simulations can be used to measure the lifetime-force dependency.



(a) *Fraction of constructs switching to the extended state as a function of force for* $N_H = 8$.



(b) *Critical force for a variation in base pairs in* N_H *and* N_A.

Figure 46: *Calculations of the critical force and fraction of strands in the extended state for parameters found using simulations. The critical force where half of the constructs switches to the extended state is higher than for the experimental values.*

5 Conclusion

A variety of complex molecular bonds have a lifetime that increases with force. To explain these observations and make it possible to apply catch bonds in humanmade materials, we tested the use of DNA to create a simple, humanmade, and tunable catch bond. The idea for these catch bonds is that there are two unbinding pathways, one for low forces and one for high forces. These DNA catch bonds where tested as a whole by measuring the lifetime-force dependency, and in detail, by checking all of the required aspects of the DNA catch bonds.

Theoretically, the requirements for a catch bond effect are satisfied for a range of constructs, with a varying number of base pairs in parts of the DNA catch bonds. The theoretical analysis suggests catch bonds can be created using DNA. Also, by varying the number of base pairs in parts of the constructs, the lifetime, critical force where the lifetime increases with force and the increase in lifetime can be tuned. Multiple catch bonds and catch bonds and slip bonds can be combined in series and in parallel to tune the lifetime/force dependency further. However, the values presented in this theoretical analysis should be interpreted as rough predictions and not as exact numerical solutions. The analysis does not include all contributions to the free energies, all possible pathways and energy landscape. Taking into account these details would change the results quantitatively.

Using experiments and simulations, we tested the catch bond effect found in theory. Each of the requirements was tested individually. Gel electrophoresis experiments with restriction enzymes show that the theoretically expected ground-state, the folded state, is the most common state for the tested construct. Performing gel electrophoresis experiments at a higher temperature shows that also the extended state forms. Together with the observation that bonds form in other experiments, these results show that the DNA catch bonds are stable and in the expected state.

Optical tweezers experiments clarify the force-extension dependency of the DNA catch bonds. The results of these experiments show an applied force can cause a change from a folded state to an extended state. So a high enough applied force can be used to change the pathway of unbinding the DNA catch bond constructs. We also observed state switching in constant force experiments. Part of the experiments did not show a switch of state, and even some force-extension measurements did not agree with theoretical force-extension curves. A possible explanation of these observations is that multiple DNA catch bonds interact with each other. More experiments must be performed to measure the critical force accurately. Improving the constructs in the way suggested in the outlook potentially increases the number of experiments where state-switching is observed.

The same optical tweezer experiments were also used to determine whether an extra, stabilizing region hybridizes. If this region hybridizes, switching from the extended state to the folded state is unlikely at finite forces. In the force-extension experiments switching to the folded state happens at a finite force, so it is unlikely that the stabilizing duplex hybridizes. More experiments are needed to check whether the region hybridizes when more time is spent at high forces. If not, the (minimal) number of base pairs in the extra region should be higher to increase the likelihood of hybridization.

Finally, the lifetime of the extended state and the folded state are compared using simulations. The simulations did not agree with the experiments. Before the results of simulations can be used to measure the difference in lifetime in the folded and the extended state, the simulation method should be improved in such a way that the simulation results agree with

the experimental results. Potentially, changing the backbone potential could improve future simulations.

The DNA catch bonds were also tested as a whole by measuring the lifetime-force dependency. The experimental methods did not give the desired results, because either many more experiments must be performed or the experimental setup must be improved. Getting enough data with optical tweezers takes a long time. This problem was absent for Centrifugal Microscopy with Nanoswitches, but the acceleration time limits the applicability of this method for constant force measurements. Simulations could not be used to find the average lifetime-force dependency because the simulational results do not agree with the experimental results.

The results in this thesis show that part of the requirements for a catch bond behavior are satisfied with DNA catch bonds. Our theoretical model suggests all requirements can be satisfied and experiments show DNA catch bonds are stable and likely satisfy the state-switching requirement. However, other aspects need further testing before can be concluded that DNA can be used to create a simple, tunable catch bond.

5.1 Outlook

Part of our initial theory, simulations and experiments support the idea that DNA can be used to create simple, humanmade, and tunable catch bonds. However, more experiments and simulations are required before conclusions can be drawn about other aspects. Both the hybridization of the stabilizing duplex and the increase in lifetime after state-switching require further tests. The tests of the lifetime-force dependency gave new inside in the mechanisms involved, but further testing is required to be conclusive about whether the DNA construct has a reversed lifetime-force dependency in a range of forces.

The hybridization of the stabilizing region can be tested by performing force-extension optical tweezer experiments, similar to the ones described in 3.3. Two aspects can be changed. First, the time spent at high forces can be increased. So instead of ramping the force up to a maximum force and immediately decrease the force, some time is spent at the maximum force. If in these experiments a forward switch but no backward switch is observed, the stabilizing region hybridizes, but only after some time. If the stabilizing region also does not hybridize over time, the number of base pairs in this region could be increased. This increases the likelihood of hybridization.

Optical tweezers, centrifugal microscopy and simulations were used to find the lifetime-force dependency of the DNA constructs. Each of these methods had its limitations. The main problems with the experiments are the time needed for the experiments and the limited force rate. As a possible way to measure the lifetime-force dependency, I suggest using parallel magnetic tweezers as reported by De Vlaminck et al. [69]. This method allows for high experimental throughput, as well as a high force-rate. Another option is Atomic Force Microscopy. This method has been used before to measure catch bond behavior in biological systems [1, 9].

Part of the performed experiments show results that do not agree with theoretical expectations. A possible cause is the interaction between multiple DNA catch bonds. In future experiments, interactions can be suppressed by slightly altering the design. Instead of using a separate region for the stabilizing region, this region is added to the stems of the hairpins. The outermost part of the hairpin stems can be chosen in such a way that it can hybridize with the other side of the hairpin and with the other strand. Figure 47 illustrates this design.

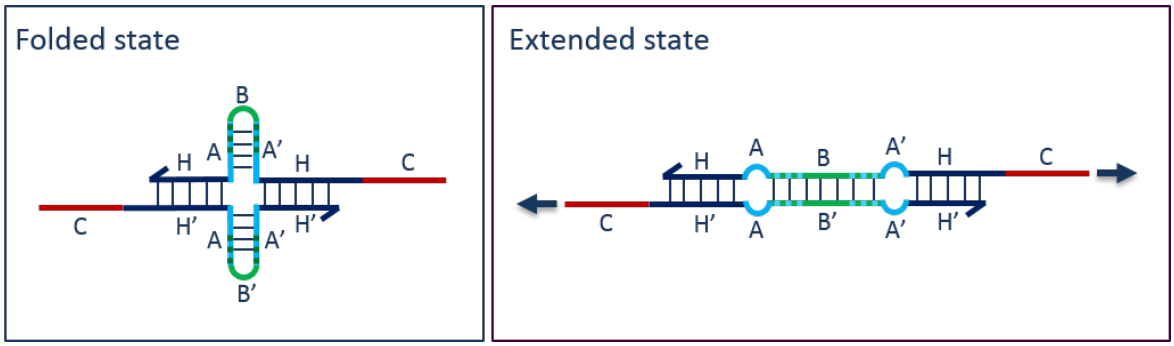


Figure 47: *Alternative design for DNA catch bonds. The stabilizing region and the stems of the hairpins are combined in one region. Due to the smaller single-stranded loops in both the folded and the extended state, the undesired interaction between constructs is reduced.*

The smaller number of unbound bases in both the folded and the extended state reduces the number of base pairs that can be formed with other complexes, reducing undesired interactions.

The results of future simulations can be improved by altering the simulational parameters. At this moment, lifetime-force simulations with oxDNA do not agree with our experimental results and the experimental results of other researchers. Altering the backbone potential used in oxDNA potentially makes this agreement better, after which the simulations can be used to make predictions about the increase in lifetime after state-switching and the lifetime-force dependency of DNA catch bonds.

5.2 Technological relevance

When the final tests of DNA catch bonds have been performed, catch bonds can be applied in materials and technologies. Here I give some suggestions for future applications.

DNA catch bonds can be used to program the lifetime-force dependency of materials. As discussed earlier, DNA catch bonds and slip bonds can be combined in series and parallel. This increases the tunability of the lifetime-force landscape. In theory, dependencies with multiple peaks and valleys can be created. Possible reported advantages for material design include enhanced ductility, toughness and durability, lower hysteresis and faster relaxation of residual strain than networks of slip bonds [16]. Some specific applications of catch bonds have been proposed in the literature. Some examples include using catch bonds as force sensors and for selective drug [17, 18].

Possibly, DNA catch bonds have alternative applications in materials. As an example, appendix E shows the theoretically expected dependency of the singlet fraction of DNA catch bond strands as a function of temperature. According to free energy calculations, the singlet fraction may depend non-monotonically on temperature. Testing other ways of using DNA catch bonds is potentially relevant for application in technologies and material optimization.

5.3 Acknowledgments

I would like to thank Kees Storm for his help and guidance with the theoretical and simulational part of this project as well as making it possible for me to work in four research groups during my master. This did not only contribute to the results in this thesis but also made my master an enjoyable experience where I learned to know various sides of research. I am particularly grateful for the assistance given by Toly Rinberg, Andrew Bergman, Vinothan Manoharan, Darren Yang, Wesley Wong, Mohsin Naqvi, Florian Wruck and Sander Tans on the experimental part of this research. Also, my colleagues at the Theory of Polymers and Soft Matter group at Eindhoven University, Manoharan lab at Harvard University, Wong lab at Harvard University and the Biophysics group at AMOLF provided me with very valuable suggestions, discussions and assistance. Finally, I would like to thank Mirjam Leunissen for proposing the research to the DNA catch bond design discussed in this thesis.

References

[1] B. T. Marshall, M. Long, J. W. Piper, T. Yago, R. P. McEver, and C. Zhu. Direct observation of catch bonds involving cell-adhesion molecules. *Nature*, 423:190–193, 2003.

[2] Brian Savage, Enrique Saldívar, and Zaverio M Ruggeri. Initiation of platelet adhesion by arrest onto fibrinogen or translocation on von willebrand factor. *Cell*, 84(2):289 – 297, 1996.

[3] Becky J. Fredrickson, Jing-Fei Dong, Larry V. McIntire, and José A. López. Shear-dependent rolling on von willebrand factor of mammalian cells expressing the platelet glycoprotein ib-ix-v complex. *Blood*, 92(10):3684–3693, 1998.

[4] MH Kroll, JD Hellums, LV McIntire, AI Schafer, and JL Moake. Platelets and shear stress. *Blood*, 88(5):1525–1541, 1996.

[5] Erik B Finger, Kamal D Puri, Ronen Alon, Michael B Lawrence, et al. Adhesion through l-selectin requires a threshold hydrodynamic shear. *Nature*, 379(6562):266, 1996.

[6] DE Brooks, J Cavanagh, D Jayroe, J Janzen, R Snoek, et al. Involvement of the mn blood group antigen in shear-enhanced hemagglutination induced by the escherichia coli f41 adhesin. *Infection and immunity*, 57(2):377–383, 1989.

[7] Zheng Jian Li, Nehal Mohamed, and Julia M. Ross. Shear stress affects the kinetics of staphylococcus aureus adhesion to collagen. *Biotechnology Progress*, 16(6):1086–1090, 2000.

[8] Wendy E. Thomas, Elena Trintchina, Manu Forero, Viola Vogel, and Evgeni V. Sokurenko. Bacterial adhesion to target cells enhanced by shear force. *Cell*, 109(7):913 – 923, 2002.

[9] Fang Kong, Andrés J García, A Paul Mould, Martin J Humphries, and Cheng Zhu. Demonstration of catch bonds between an integrin and its ligand. *The Journal of cell biology*, 185(7):1275–1284, 2009.

[10] Derek L. Huang, Nicolas A. Bax, Craig D. Buckley, William I. Weis, and Alexander R. Dunn. Vinculin forms a directionally asymmetric catch bond with f-actin. *Science*, 357(6352):703–706, 2017.

[11] Sabyasachi Rakshit, Yunxiang Zhang, Kristine Manibog, Omer Shafraz, and Sanjeevi Sivasankar. Ideal, catch, and slip bonds in cadherin adhesion. *Proceedings of the National Academy of Sciences*, 109(46):18815–18820, 2012.

[12] Wendy E. Thomas, Viola Vogel, and Evgeni Sokurenko. Biophysics of catch bonds. *Annual Review of Biophysics*, 37(1):399–416, 2008. PMID: 18573088.

[13] W Benjamin Rogers and Vinothan N Manoharan. Programming colloidal phase transitions with dna strand displacement. *Science*, 347(6222):639–642, 2015.

[14] Aled D. Roberts, Xu Li, and Haifei Zhang. Porous carbon spheres and monoliths: morphology control, pore size tuning and their applications as li-ion battery anode materials. *Chem. Soc. Rev.*, 43:4341–4356, 2014.

[15] Xiaoshan Fan, Jing Yang Chung, Yong Xiang Lim, Zibiao Li, and Xian Jun Loh. Review of adaptive programmable materials and their bioapplications. *ACS Applied Materials & Interfaces*, 8(49):33351–33370, 2016. PMID: 27960431.

[16] Balaji V.S. Iyer, Victor V. Yashin, and Anna C. Balazs. Harnessing biomimetic catch bonds to create mechanically robust nanoparticle networks. *Polymer*, 69:310 – 320, 2015. Self-Healing Polymers.

[17] Manu Forero, Wendy E. Thomas, Clint Bland, Lina M. Nilsson, Evgeni V. Sokurenko, and Viola Vogel. A catch-bond based nanoadhesive sensitive to shear stress. *Nano Letters*, 4(9):1593–1597, 2004.

[18] Elizaveta A Novikova and Cornelis Storm. Contractile fibers and catch-bond clusters: a biological force sensor? *Biophysical journal*, 105(6):1336–45, 2013.

[19] Kary B Mullis and Fred A Faloona. [21] specific synthesis of dna in vitro via a polymerase-catalyzed chain reaction. *Methods in enzymology*, 155:335–350, 1987.

[20] Matthias Rief, Hauke Clausen-Schaumann, and Hermann E Gaub. Sequence-dependent mechanics of single dna molecules. *Nature Structural & Molecular Biology*, 6(4):346–349, 1999.

[21] Dominik Ho, Julia L Zimmermann, Florian A Dehmelt, Uta Steinbach, Matthias Erdmann, Philip Severin, Katja Falter, and Hermann E Gaub. Force-driven separation of short double-stranded dna. *Biophysical journal*, 97(12):3158–3167, 2009.

[22] John SantaLucia Jr., Hatim T. Allawi, , and P. Ananda Seneviratne. Improved Nearest-Neighbor Parameters for Predicting DNA Duplex Stability. *Biochemistry*, 35(11):3555–3562, 1996. PMID: 8639506.

[23] John SantaLucia. A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proceedings of the National Academy of Sciences of the United States of America*, 95(4):1460–1465, 1998.

[24] John SantaLucia and Donald Hicks. The thermodynamics of DNA structural motifs. *Annual review of biophysics and biomolecular structure*, 33:415–40, 2004.

[25] W F Bailey and A S Monahan. Statistical effects and the evaluation of entropy differences in equilibrium processes. symmetry corrections and entropy of mixing. *Journal of Chemical Education*, 55(8):489–493, 1978.

[26] Manoel Manghi, Nicolas Destainville, and John Palmeri. Mesoscopic models for dna stretching under force: New results and comparison with experiments. *The European Physical Journal E*, 35(10):110, Oct 2012.

[27] Majid Mosayebi, Ard A Louis, Jonathan PK Doye, and Thomas E Ouldridge. Force-induced rupture of a dna duplex: from fundamentals to force sensors. *ACS nano*, 9(12):11993–12003, 2015.

[28] Torsten Strunz, Krisztina Oroszlan, Rolf Schäfer, and Hans-Joachim Güntherodt. Dynamic force spectroscopy of single dna molecules. *Proceedings of the National Academy of Sciences*, 96(20):11277–11282, 1999.

[29] K Hatch, C Danilowicz, V Coljee, and M Prentiss. Demonstration that the shear force required to separate short double-stranded dna does not increase significantly with sequence length for sequences longer than 25 base pairs. *Physical Review E*, 78(1):011920, 2008.

[30] Pierre-Gilles de Gennes. Maximum pull out force on {DNA} hybrids. *Comptes Rendus de l'Académie des Sciences - Series {IV} - Physics*, 2(10):1505 – 1508, 2001.

[31] 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. *The Journal of Chemical Physics*, 137(13):135101, 2012.

[32] Evgeni V Sokurenko, Viola Vogel, and Wendy E Thomas. Catch-bond mechanism of force-enhanced adhesion: counterintuitive, elusive, but... widespread? *Cell host & microbe*, 4(4):314–323, 2008.

[33] Masashi Suzuki, Naoto Yagi, and John T Finch. Role of base-backbone and base-base interactions in alternating dna conformations. *FEBS letters*, 379(2):148–152, 1996.

[34] L. J. McBride and M. H. Caruthers. An investigation of several deoxynucleoside phosphoramidites useful for synthesizing deoxyoligonucleotides. *Tetrahedron Letters*, 24(3):135101, 1982.

[35] Brian R. Wolfe, Nicholas J. Porubsky, Joseph N. Zadeh, Robert M. Dirks, and Niles A. Pierce. Constrained multistate sequence design for nucleic acid reaction pathway engineering. *Journal of the American Chemical Society*, 139(8):3134–3144, 2017. PMID: 28191938.

[36] Rudolphus G. H. van Leeuwen. Protein folding and translocation: Single-molecule investigationsl. *U. Leiden*, 2006.

[37] Karel Svoboda, Christoph F. Schmidt, Bruce J. Schnapp, and Steven M. Block. Direct observation of kinesin stepping by optical trapping interferometry. *Nature*, 365(6448):721–727, 10 1993.

[38] K Svoboda and S M Block. Biological applications of optical forces. *Annual Review of Biophysics and Biomolecular Structure*, 23(1):247–285, 1994. PMID: 7919782.

[39] 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.

[40] Darren Yang, Andrew Ward, Ken Halvorsen, and Wesley P Wong. Multiplexed single-molecule force spectroscopy using a centrifuge. *Nature communications*, 7, 2016.

[41] Modesto Orozco, Alberto Pérez, Agnes Noy, and F Javier Luque. Theoretical methods for the simulation of nucleic acids. *Chemical Society reviews*, 32:350–64, 12 2003.

[42] David Swigon. *The Mathematics of DNA Sturcture, Mechanics, and Dynamics*, volume 150, pages 293–320. Springer Science, 01 1970.

[43] Alberto Pérez, F. Javier Luque, and Modesto Orozco. Dynamics of b-dna on the microsecond time scale. *Journal of the American Chemical Society*, 129(47):14739–14745, 2007. PMID: 17985896.

[44] Alex Morriss-Andrews, Joerg Rottler, and Steven S. Plotkin. A systematically coarse-grained model for dna and its predictions for persistence length, stacking, twist, and chirality. *The Journal of Chemical Physics*, 132(3):035105, 2010.

[45] William B. Sherman and Nadrian C. Seeman. A precisely controlled dna biped walking device. *Nano Letters*, 4(7):1203–1207, 2004.

[46] Thomas E. Ouldridge, Ard A. Louis, and Jonathan P. K. Doye. Dna nanotweezers studied with a coarse-grained model of dna. *Phys. Rev. Lett.*, 104:178101, Apr 2010.

[47] Flavio Romano, Alexander Hudson, Jonathan P. K. Doye, Thomas E. Ouldridge, and Ard A. Louis. The effect of topology on the structure and free energy landscape of dna kissing complexes. *The Journal of Chemical Physics*, 136(21):215102, 2012.

[48] Petr Šulc, Thomas E. Ouldridge, Flavio Romano, Jonathan P. K. Doye, and Ard A. Louis. Simulating a burnt-bridges dna motor with a coarse-grained dna model. *Natural Computing*, 13(4):535–547, Dec 2014.

[49] Flavio Romano, Debayan Chakraborty, Jonathan P. K. Doye, Thomas E. Ouldridge, and Ard A. Louis. Coarse-grained simulations of dna overstretching. *The Journal of Chemical Physics*, 138(8):085101, 2013.

[50] 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. PMID: 26575598.

[51] Charles Jaffé, Shane D. Ross, Martin W. Lo, Jerrold Marsden, David Farrelly, and T. Uzer. Statistical theory of asteroid escape rates. *Phys. Rev. Lett.*, 89:011101, Jun 2002.

[52] Christoph Dellago and Peter G. Bolhuis. Transition path sampling and other advanced simulation techniques for rare events. *Advances in Polymer Science*, pages 167–233, 2009.

[53] Peter Hänggi, Peter Talkner, and Michal Borkovec. Reaction-rate theory: fifty years after kramers. *Rev. Mod. Phys.*, 62:251–341, Apr 1990.

[54] Titus S. Van Erp. *Dynamical Rare Event Simulation Techniques for Equilibrium and Nonequilibrium Systems*, chapter 2, pages 27–60. Wiley-Blackwell, 2012.

[55] Titus S Van Erp. Dynamical rare event simulation techniques for equilibrium and nonequilibrium systems. *Advances in Chemical Physics*, 151:27, 2012.

[56] Peter Májek and Ron Elber. Milestoning without a reaction coordinate. *Journal of chemical theory and computation*, 6(6):1805–1817, 2010.

[57] Nina Singhal, Christopher D Snow, and Vijay S Pande. Using path sampling to build better markovian state models: predicting the folding rate and mechanism of a tryptophan zipper beta hairpin. *The Journal of chemical physics*, 121(1):415–425, 2004.

[58] Titus S Van Erp and Peter G Bolhuis. Elaborating transition interface sampling methods. *Journal of computational Physics*, 205(1):157–181, 2005.

[59] Daniele Moroni, Titus S van Erp, and Peter G Bolhuis. Investigating rare events by transition interface sampling. *Physica A: Statistical Mechanics and its Applications*, 340(1):395–401, 2004.

[60] Titus S van Erp, Daniele Moroni, and Peter G Bolhuis. A novel path sampling method for the calculation of rate constants. *The Journal of chemical physics*, 118(17):7762–7774, 2003.

[61] Daniele Moroni, Peter G Bolhuis, and Titus S van Erp. Rate constants for diffusive processes by partial path sampling. *The Journal of chemical physics*, 120(9):4055–4065, 2004.

[62] Anthony MA West, Ron Elber, and David Shalloway. Extending molecular dynamics time scales with milestoning: example of complex kinetics in a solvated peptide. *The Journal of chemical physics*, 126(14):04B608, 2007.

[63] Anton K Faradjian and Ron Elber. Computing time scales from reaction coordinates by milestoning. *The Journal of chemical physics*, 120(23):10880–10889, 2004.

[64] Eric Vanden-Eijnden, Maddalena Venturoli, Giovanni Ciccotti, and Ron Elber. On the assumptions underlying milestoning. *The Journal of chemical physics*, 129(17):174102, 2008.

[65] Rosalind J Allen, Chantal Valeriani, and Pieter Rein ten Wolde. Forward flux sampling for rare event simulations. *Journal of physics: Condensed matter*, 21(46):463102, 2009.

[66] Rosalind J Allen, Daan Frenkel, and Pieter Rein ten Wolde. Forward flux sampling-type schemes for simulating rare events: Efficiency analysis. *The Journal of chemical physics*, 124(19):194111, 2006.

[67] Peter Reimann, GJ Schmid, and P Hänggi. Universal equivalence of mean first-passage time and kramers rate. *Physical Review E*, 60(1):R1, 1999.

[68] Thomas E Ouldridge, Ard A Louis, and Jonathan PK Doye. Structural, mechanical, and thermodynamic properties of a coarse-grained dna model. *The Journal of chemical physics*, 134(8):02B627, 2011.

[69] Iwijn De Vlaminck, Thomas Henighan, Marijn T.J. van Loenhout, Indriati Pfeiffer, Julius Huijts, Jacob W. J. Kerssemakers, Allard J. Katan, Anja van Langen-Suurling, Emile van der Drift, Claire Wyman, and Cees Dekker. Highly parallel magnetic tweezers by targeted dna tethering. *Nano Letters*, 11(12):5489–5493, 2011. PMID: 22017420.

A Sequences

Gel electrophoresis experiments and centrifugal microscopy experiments:

$$
\begin{aligned}
Seq\ 1 &: GTTGAAGTCGAAGTCCTGATAGTGA - TTATAATTACAA - AGAGCGGA - \\
&\quad TTGTAATTATAA - GCCTAATAATCAATCTGTAAGCAAG - TTTTAACGT \\
&\quad AAGGAAGAAGAAGTAATAATAAGAGTCGGAATAGATGAAGAGAGCCAG \\
&\quad AGAGAAG \\
Seq\ 2 &: CTTGCTTACAGATTGATTATTAGGC - AATATTAATGTT - TCCGCTCT - \\
&\quad AACATTAATATT - TCACTATCAGGACTTCGACTTCAAC - TTTTTAGGA \\
&\quad ATACCACATTCAACTAATGCAGATACATAACGCCAAAAAGGAATTACGA \\
&\quad GGCATAG \\
Seq\ 3 &: CGGTCGGATC - TTATAATTACAA - AGAGCGGA - TTGTAATTATAA - \\
&\quad CCAAGCACAA - TTTTAACGTAAGGAAGAAGAAGTAATAATAAGAGTCGG \\
&\quad AATAGATGAAGAGAGCCAGAGAGAAG \\
Seq\ 4 &: TTGTGCTTGG - AATATTAATGTT - TCCGCTCT - AACATTAATATT - \\
&\quad GATCCGACCG - TTTTTAGGAATACCACATTCAACTAATGCAGATACATA \\
&\quad ACGCCAAAAAGGAATTACGAGGCATAG \\
Seq\ 5 &: TCGGATC - TTATAATTACAA - AGAGCGGA - TTGTAATTATAA - CCAAG \\
&\quad CA - TTTTAACGTAAGGAAGAAGAAGTAATAATAAGAGTCGGAATAGAT \\
&\quad GAAGAGAGCCAGAGAGAAG \\
Seq\ 6 &: TGCTCTG - AATATTAATGTT - TCCGCTCT - AACATTAATATT - GATCC \\
&\quad GA - TTTTTAGGAATACCACATTCAACTAATGCAGATACATAACGCCAA \\
&\quad AAGGAATTACGAGGCATAG
\end{aligned}
\tag{73}
$$

Optical tweezer experiments:

$$
\begin{aligned}
Seq\ 1: &\ AGTCAATCTAGAGAGCCCTG - AAAAA - CATAATAACC - T - GTGAGT \\
&\ GAGGGAGTG - GCCTCGTCCG - CACTCCCTCACTCAC - T - CATACAT \\
&\ CAC \\
Seq\ 2: &\ AGTCAATCTAGAGAGCCCTG - AAAAA - GTGATGTATG - T - CCACAA \\
&\ ACCACCACA - CGGACGAGGC - TGTGGTGGTTTGTGG - T - GGTTATT \\
&\ ATG \\
Seq\ 3: &\ AGTCAATCTAGAGAGCCCTG - AAAAA - CGAATATAAACC - T - CAAC \\
&\ ACAACCACCTTCCC - GCCACCGAGGGA - GGGAAGGTGGTTGTGTTG- \\
&\ T - CTCCACTCCAAC \\
Seq\ 4: &\ AGTCAATCTAGAGAGCCCTG - AAAAA - GTTGGAGTGGAG - T - GTTT \\
&\ AGTAGTAGAGAGGA - TCCCTCGGTGGC - TCCTCTCTACTACTAAAC- \\
&\ T - GGTTTATATTCG
\end{aligned}
\tag{74}
$$

B oxDNA protocol

The simulations with oxDNA are performed in five (automated) steps. First, several parameters for the simulation should be defined in four files. First of all, the desired sequences are defined in *seq.txt*. Either the exact sequence can be defined (using A, T, G and C) or an undefined sequence, where only complementary regions are indicated. When an unspecific sequence is chosen, every base is assumed different from all other bases and only complementary to its counterpart. This corresponds to the (unrealistic) situation of a perfectly designed sequence without off-target interactions. Complementary regions are indicated with (arbitrary) complementary letters, which can be chosen and defined at will. In the same way, self-complementarity (binding to the strand the bases are part of) or non-complementarity (binding to nothing) is defined with letters. Each line in the sequence file corresponds to one sequence. For example the sequence file:

BBBBBZZZZZ
ZZZZZDDDDD

results in a duplex with B and D bound with two single-stranded tails (Z is defined to by non-complementary). The sequence file:

NNNNNZZZZZOOOOO

results in a hairpin. Also, letters can be chosen that correspond to complementary regions, but not bound in the initial state. New combinations of letters can be defined in a *run.sh*. *relaxing_condition.txt*, contains the minimal number of base pairs formed before a structure is considered relaxed.

In the third input file *op.txt*, can be defined what should be tracked during the simulation. In our case, this is the distance between complementary bases. With this data, we can check whether, at a moment in time, bases are hybridized or not. The data provided in this file can be used as an order parameter for Stairway Sampling. Multiple order parameters can be defined and tracked at the same time. For example the order parameter file:

```
{
order_parameter = bond
name = bond1
pair1 = 0, 19
}
```

tracks the distance between base 0 and 19. The value of *bond*1 will be 1 if the bases are at a base-pairing distance, in the correct orientation and 0 when they are in any other position.

Finally, some important parameters that change regularly can be defined in *run.sh*. These include for example the used temperature, salt concentration and the desired number of simulations.

After providing these input files, you can start the simulation. The program starts with the generating an initial configuration. This configuration is based on the provided sequence file. Each strand is placed with a random (but not overlapping) location and orientation in simulation space.

Next, this generated configuration is used to create ten initial configurations for the simulation. This relaxation protocol starts with bringing complementary bases that should be bound in the initial state close to each other using an MC simulation. The backbone potential

is replaced by an (unrealistic) potential that does not have a corresponding force that diverges with distance. An additional, infinite potential is added for every base that should be bound in the initial state. This is a potential that scales linearly with distance with a minimum at the distance where bases form base pairs. The relaxation is completed by running an MD simulation until the stopping criterion, as defined in *relaxing_condition.txt*, is reached. The end configuration is then used in another MC simulation followed by an MD simulation to create a second starting configuration. This process is repeated until ten starting configurations are formed.

Now the Stairway Sampling protocol is started. One of the starting configurations is selected at random, and a short simulation is used to create a variation on this starting configuration. Small pieces of tracks are simulated, which the program will later combine to get valuable information about the process. During these simulations, the order of unbinding of individual bases is tracked, so the path of unbinding can be followed. After finishing all simulations, the results are summarized and combined to give the average unbinding pathways and times. Finally, figures of the results are automatically generated.

So to summarize, input options must be provided in *seq.txt*, *op.txt*, *relaxation_condition.txt* and *run.sh*. This information is used to generate a starting configuration, which is relaxed. Then Langevin Dynamics simulations using oxDNA are combined with Forward Stairway Sampling to generate data. Finally, this data is combined to find the unbinding times and paths.

C Gel electrophoresis pictures

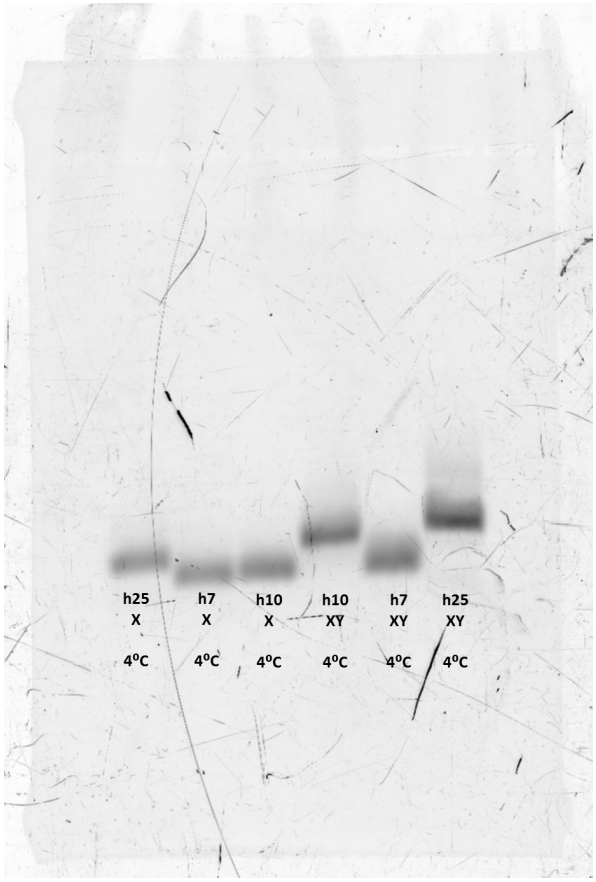


Figure 48: *Picture of gel after 60 minutes of gel electrophoresis at 4C. The used gel is a 12% PAGE gel and the used buffer is NEBuffer Cutsmart.*

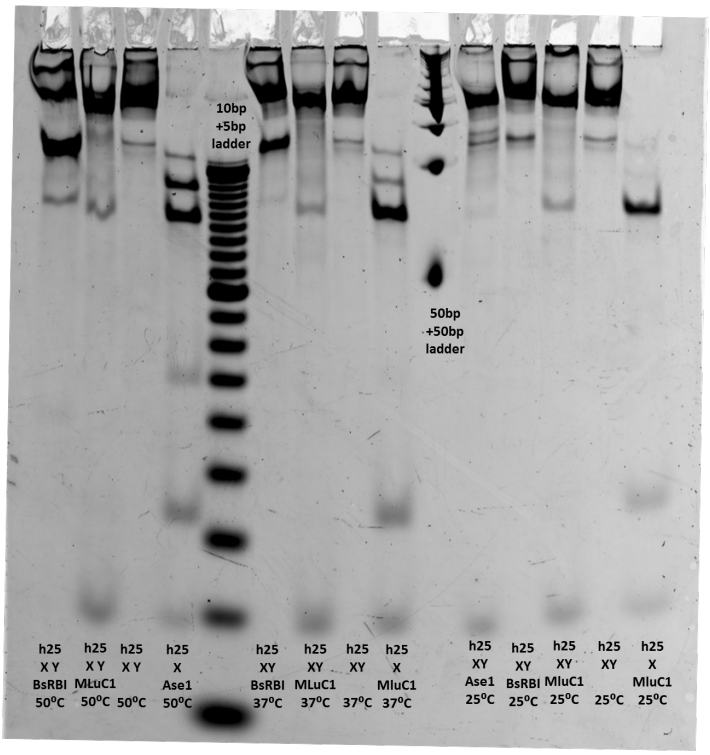


Figure 49: *Picture of gel after gel electrophoresis at 25C. The used gel is a 12% PAGE gel and the used buffer is NEBuffer Cutsmart. Indicated temperatures are the incubation temperatures.*

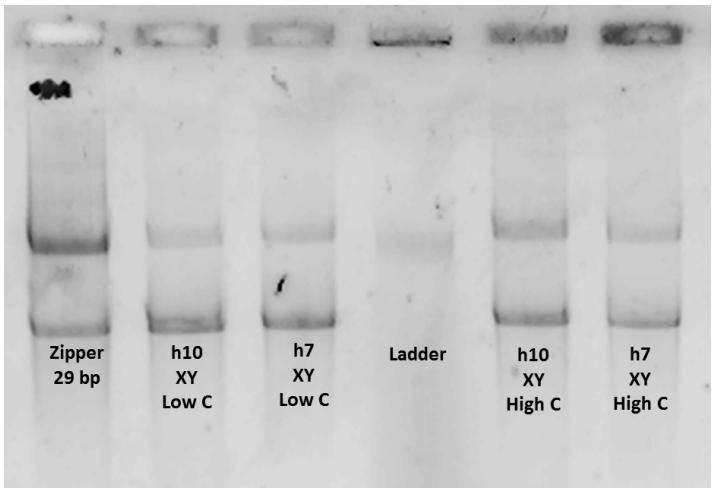


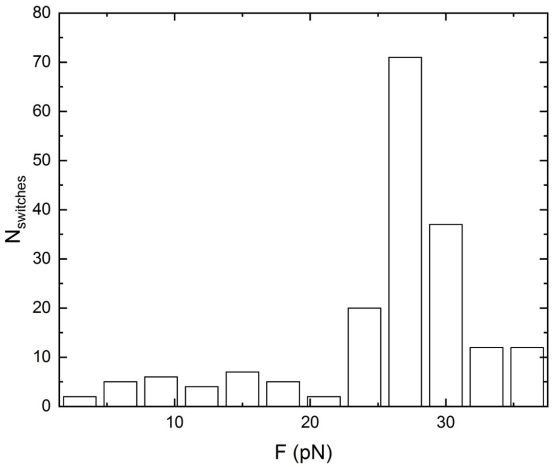
Figure 50: *Picture of gel after gel electrophoresis at 4C. The used gel is a 0.7% agarose gel and the used buffer is NEBuffer Cutsmart. Indicated temperatures are the incubation temperatures. Strand X is the final strand added to close the nanoswitches. Low C: Concentration strand X 1.25x concentration of other DNA strands, High C: Concentration strand X 2x concentration of other DNA strands)Varying the concentration of this strand does not change the percentage of nanoswitches closed significantly.*

D Lifetime-force rate dependency

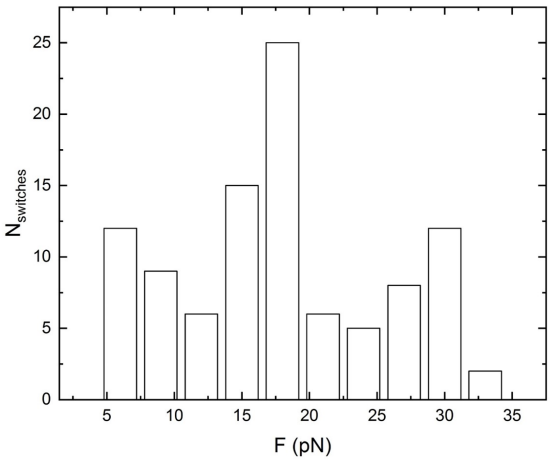
This section presents additional results of centrifugal microscopy experiments at three force rates. This data can be used to estimate what happens when the force rate is increased. The amount of data presented here is minimal, so more experiments must be performed before any conclusions can be drawn.

Figure 51a shows the results of the experiments with the DNA catch bond construct with $N_H = 10$ with force-rate between 1 and 2 pN at 15°C. If we assume the distribution to be Gaussian, we find an average and standard deviation of 27.4 ± 1.9 pN. An unusual effect occurs when the force rate is increased. Figure 51b shows a histogram of the breaking forces of the DNA catch bond with $N_H = 10$ for a force-rate of $3 - 4$ pN/s. Some of the constructs still dissociate at the same force, but also a high number of constructs break at lower forces. The same thing is observed when the force-rate is increased even more. Although more data should be gathered before real conclusions can be drawn, the data so far suggests that the lifetime-force rate dependency is noteworthy.

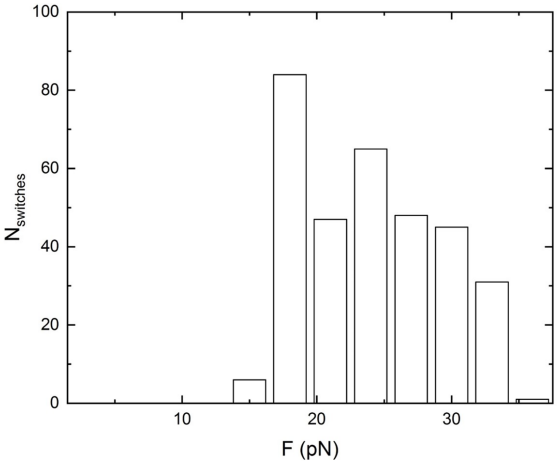
For regular slip-bonds, the average unbinding force increases with increasing force rate. For the DNA catch bonds, the data suggests that the average force could be decreasing with increasing force rate. This effect can be explained by a switch of state of the DNA catch bonds. At a low force rate, the A/A'-regions are loaded for a long enough time such that most of the constructs switch from the folded state to the extended state. However, when a higher force rate is used, this switch takes longer than the time needed to accelerate to the force where the handles break quickly. When the maximum force is reached in a force rate experiment, the force stays constant for some time before the centrifuge starts slowing down. During the time at this constant force, also dissociations are observed.



(a) *Combination of data from eight* $N_H = 10$ *experiments at* $10°C$ *with an acceleration rate of 1-2 pN/s.*



(b) *Combination of data from four* $N_H = 10$ *experiments at* $10°C$ *with an acceleration rate of 3-4 pN/s.*



(c) *Combination of data from eleven* $N_H = 10$ *experiments at* $10°C$ *with an acceleration rate of 5-6 pN/s.*

Figure 51: *Histograms from experiments with DNA catch bonds with* $N_H = 10$*. The figures compare the distribution in average breaking forces varying the force rate at* $10°C$*.*

E Alternative use of DNA catch bonds

In the previous sections, we discussed the catch bond behavior of the DNA catch bonds. In this appendix, I suggest another possible application of DNA catch bonds. From the results of the gel electrophoresis experiments, we already saw that the number of constructs in the extended state increases with temperature. This increase is caused by the larger entropy in the extended state than in the folded state. Here, I make a prediction of the dependence of the singlet-fraction of free DNA strands using free energy calculations. Figure 52 shows an example of a possible dependency of the singlet fraction on temperature. According to these calculations, the singlet fraction can be non-monotonic for DNA catch bonds.

The theoretical singlet-fraction dependency first agrees with the situation where only the folded state can be formed (red line) and then switches to the situation where only the extended state can be formed (blue line). The red line is calculated by replacing the B-regions with poly-T strands and the blue line by replacing the A-regions with poly-T strands. The melting temperature of the extended state is higher than the melting temperature of the folded state. So when the temperature is increased from the temperature where most constructs are in the folded state to the temperature where most constructs are extended, the singlet fraction can decrease. So these simple calculations suggest that DNA catch bonds can be used to create non-monotonic singlet-fraction-temperature dependencies. This could be used to tune the temperature dependency of materials and technology.

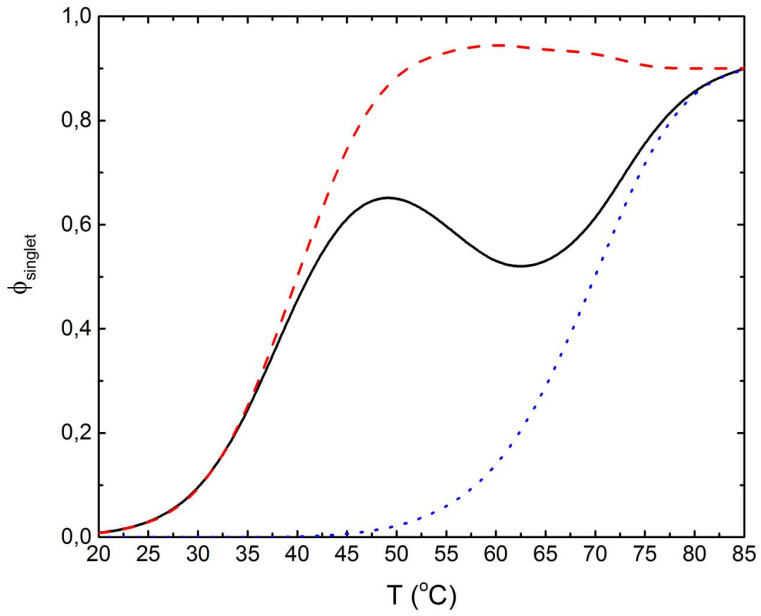


Figure 52: *Free energy calculations of singlet-fractions for DNA catch bonds as a function of temperature. The black line gives the dependency for the full construct, the red dashed line for constructs with one of the B-regions replaced by poly-T and the blue dotted line for constructs with the A-regions replaced with poly-T.*