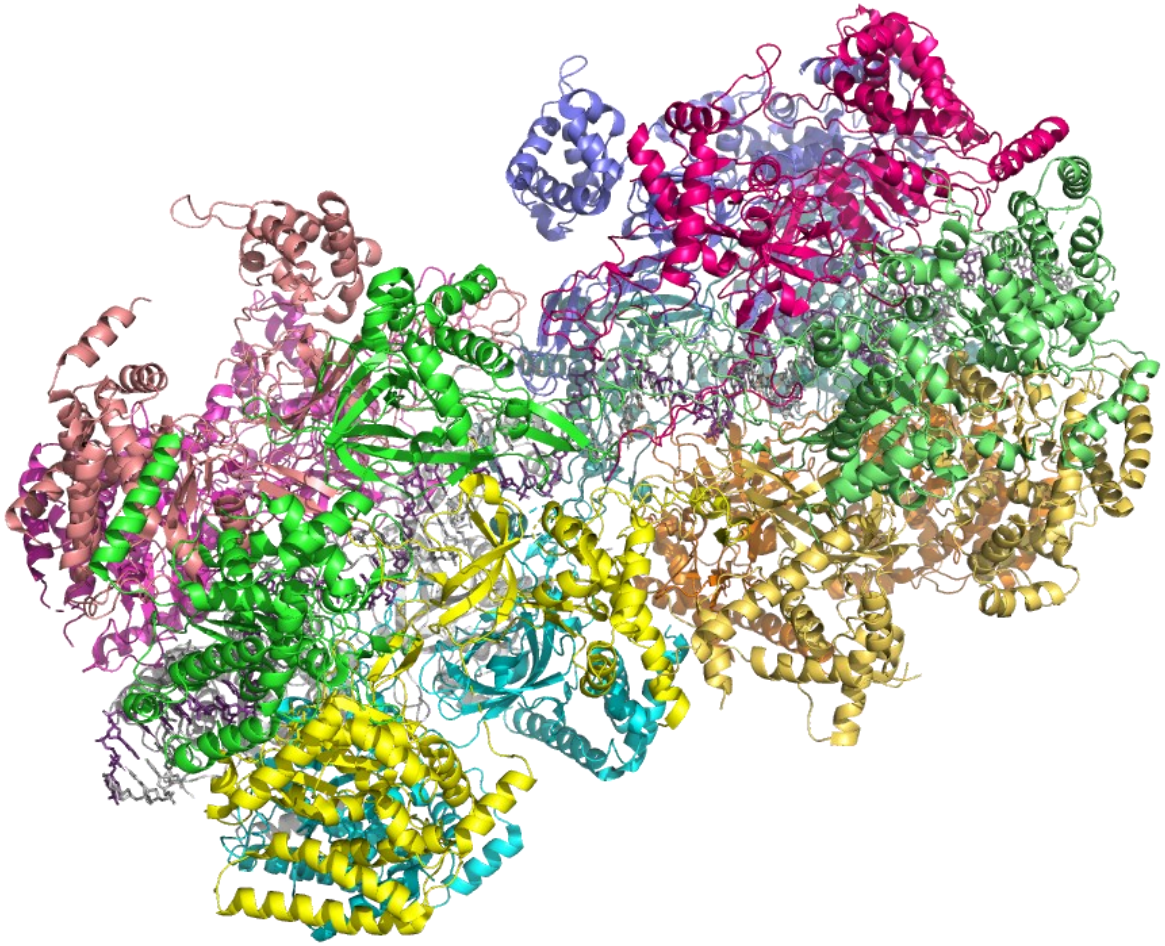
Synthesis and characterization of DNA constructs for *in vitro* single-molecule studies of eukaryotic DNA replication termination



Vincent P. E. Kruit

Bachelor End Project for BSc. in Nanobiology

Supervisors

MSc. Daniel Ramírez Montero

Prof. Dr. Nynke Dekker

Performed at

Nynke Dekker lab

Department of Bionanoscience

Delft University of Technology

Contents

Introduction

DNA replication is an essential process in the life cycle of cells as it ensures that genetic information is accurately passed down from generation to generation. Replication errors and replicative stress introduce mutations which can cause a broad spectrum of diseases like for example cancer and neurodegeneration[1]. The process of DNA replication can be divided into three subsequent stages: initiation, elongation, and termination. Although still an area of active investigation, relatively a lot is known about DNA replication initiation and elongation. In contrast, little is known about the termination stage, which will be the focus of this study.

DNA replication is the result of two concerted processes, namely unwinding of the parental DNA duplex and synthesis of the complementary strands. This process is catalyzed by the replisome, a dynamic protein complex, which is centrally organized by the Cdc45-Mcm2-7-GINS (CMG) helicase. The catalytic core of the CMG helicase is composed of six minichromosome maintenance (Mcm2-7) proteins, which form a heterohexameric ring. The mechanism by which the strands are separated is still an area of active investigation[2].

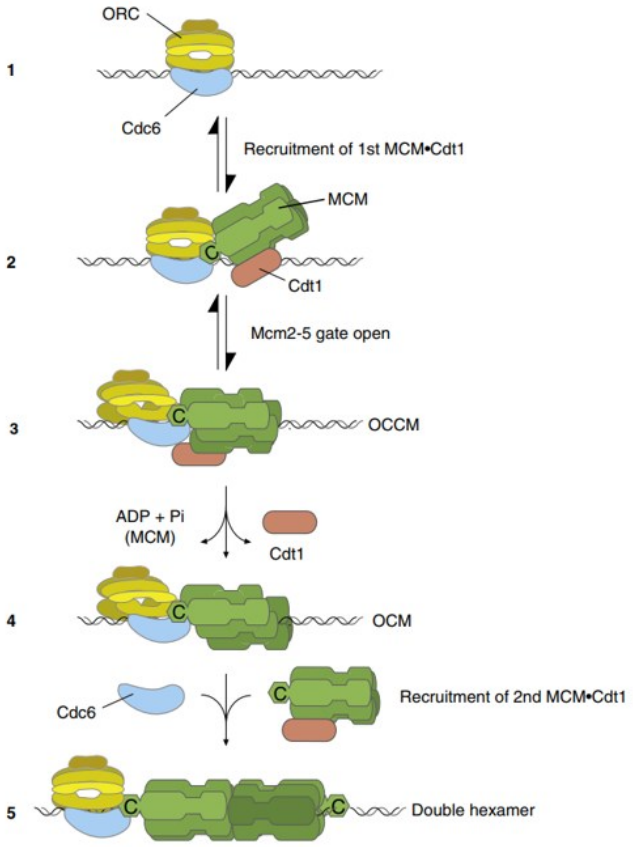


***Figure 1**. During G1-phase, double hexamers are loaded onto origins of replication by the combined action of ORC, Cdc6, and Cdt1 which recruit and load two MCMs in opposite orientations. (**1**) ORC-Cdc6 is halted by origin recognition. (**2**) The MCM-Cdt1 heptamer is recruited by the ORC-Cdc6 complex. (**3**) The recruited MCM-Cdt1 complex changes its conformation by opening of the Mcm2-5 gate to encircle dsDNA, forming an OCCM complex together with ORC and Cdc6. (**4**) ATP hydrolysis and Cdt1 release are then required to close the Mcm2-5 gate. (**5**) Another MCM is subsequently loaded in a similar manner in opposite orientation. Figure adapted from Deegan & Diffley, 2016.*

During the G1 phase of the cell cycle, DNA licensing takes place in which two Mcm2-7 (MCM) complexes are loaded as double hexamers (DH) onto double-stranded DNA (**Fig. 1**). In *S. cerevisiae,* this process occurs at defined DNA sequences known as the autonomously replicating sequence 1 (ARS1)[4,5]. The loading process is carried out by the combined action of the origin recognition complex (ORC), Cell division cycle 10-dependent transcription 1 (Cdt1), and Cell division cycle 6 (Cdc6)[6,7]. During the subsequent S phase of the cell cycle, the loaded MCM helicases are activated by the recruitment of helicase activating factors, resulting in assembly of functional CMG helicases[8].

However, only few loaded MCM helicases are activated during the S phase. This phenomenon is commonly referred to as the MCM-paradox and it is thought that some of these excess MCMs act as so-called 'dormant origins of replication' which function as a backup and may be activated when the cell experiences replicative stress, maintaining genome integrity by rescuing stalled replication forks[9,10]**.** These 'dormant origins of replication' can pose a problem for the cell during replication termination, when two converging replisomes meet.

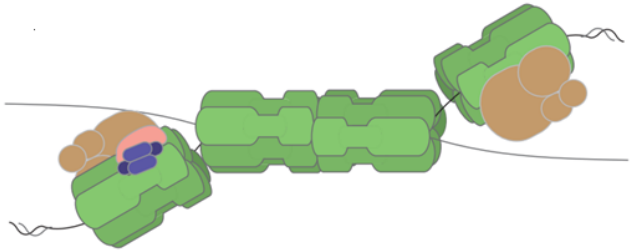


Figure 2*. Schematic representation of two converging CMG complexes obstructed by an inactive DH, during replication termination. Figure taken from Hill et al., 2020.*

During replication termination, two converging CMG complexes normally bypass each other after which the complexes are polyubiquitylated and unloaded from the chromatin[12,13]**.** However, CMG complexes may encounter inactive MCM helicases whilst translocating. CMG cannot overcome roadblocks surrounding double-stranded DNA, such as MCM helicases, but it was shown that CMG is able to push the inactive MCMs[8]. Two converging replisomes may encounter the same inactive MCMs, which would stall both complexes for they are pushing the same MCMs in opposite directions (**Fig. 2**). That inactive MCMs may obstruct CMG convergence during replication termination was not known until very recently.

Deegan et al. (2019) found that an auxiliary DNA helicase, Pif1, aids in the removal of inactive MCMs during replication termination, promoting fork convergence. Pif1 was independently found to aid in this process, but its interaction with inactive MCM helicase was not studied in detail.

Already quite a bit is known about the process of replication termination, however most studies elucidate this mechanism via a top-down approach by cellular and cell extract studies[15,12,16]. The goal of this study is to synthesise a DNA construct which allows others to directly visualise the interactions of converging CMG complexes, inactive MCMs, and Pif1 *in vitro* on the single-molecule level, using the purified yeast system[17].

Some of the outstanding questions this study will help to elucidate are whether the inactive MCMs stop CMG convergence, which has not been observed directly; by which mechanisms Pif1 aids in the removal of DHs; whether Pif1 is sufficient to remove the inactive MCMs; and lastly, whether Pif1 translocates with CMG or whether it is specifically recruited to stalled helicases.

The main focus will be the synthesis and characterization of a DNA construct for *in vitro* single-molecule studies of replication termination (**Fig. 3**). Using this construct, we hope to directly observe the obstruction of converging CMGs by inactive MCMs, which has not been done before and will give us unprecedented mechanistic insight into the eukaryotic DNA replication termination mechanism. Moreover, the constructs can be used to study the pushing of inactive MCM helicases by translocating CMGs and allows us to study the role and interactions of auxiliary DNA helicases in resolving the stalled structure. Correlative dual beam optical trapping combined with confocal microscopy will be used to

biophysically characterize the DNA construct and its interaction with various fluorescently labelled proteins, including dCas9 and MCM helicases, during *in vitro* single-molecule experiments. Moreover, biochemical techniques such as DNA sequencing and gel electrophoresis are used to biochemically characterize the construct in detail.

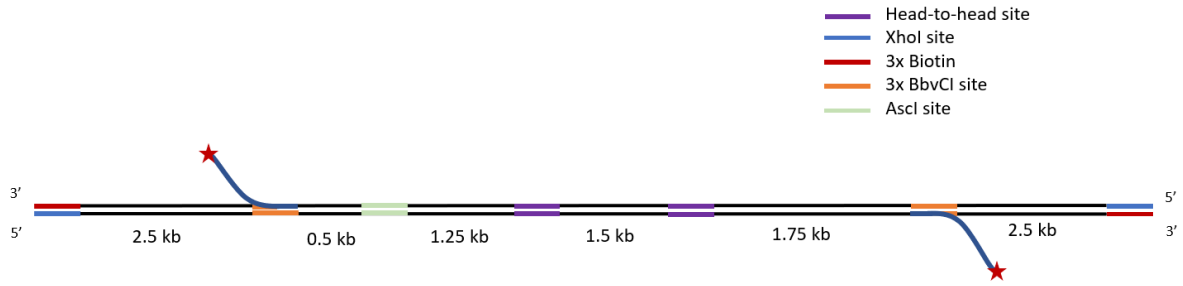


***Figure 3**. Schematic overview of the synthesized DNA construct. (**Purple**) Artificial head-to-head (HTH) origins of replication sites. (**Orange**) Fork insertion sites. (**Red**) Biotinylated nucleotides. (**Lime**) AscI restriction site. **(Blue**) XhoI restriction site.*

The synthesized DNA construct features two artificial Head-to-Head (HTH) origins of replications, each containing two sites with enhanced binding affinity for the ORC complex, compared to the naturally occurring ARS1 sites[4]. At the HTH sites, DHs will be loaded. Then, to study the interaction of CMG and Pif1 with these DHs, purified CMG complexes will be used which need a ssDNA 3' overhang to bind and start translocating in the 3' to 5' direction, fuelled by the hydrolysis of ATP. To this end, two fluorescently labelled 3' polyT tails are incorporated into the construct. These fluorescently labelled forks simultaneously function as a nanoruler making it possible to determine distances between other fluorescently labelled proteins with high precision.

Additionally, the construct features an off-center AscI restriction site which can be used to determine the fork incorporation efficiency of either fork. Furthermore, the construct features off-center dCas9 binding sites, to which fluorescently labelled dCas9 can bind, making it possible to determine the orientation of single optically trapped construct molecules in the single-molecule setup. The construct is made from a plasmid which is linearized by XhoI restriction enzymes, the resulting sticky ends are biotinylated to be able to optically trap the construct during single-molecule experiments using streptavidin-coated polystyrene beads.

In addition to construct synthesis and characterization, a bacterial overexpression plasmid containing a single-solvent-exposed-cysteine mutant of Pif1 is generated. This mutant version of Pif1 will allow us to site-specifically fluorescently label Pif1 using cysteine-maleimide chemistry to be used in single-molecule experiments. Following mutagenesis, the entire open reading frame of the mutated Pif1 sequence is sequenced for both strands to characterize it and confirm successful mutagenesis.

Theoretical background

T.1 Eukaryotic DNA replication

DNA replication is an essential process that ensures genetic information is faithfully transmitted from generation to generation. This process is carried out by a dynamic protein complex known as the replisome, which seems to be largely conserved among eukaryotes[18]. The replisome duplicates the parental DNA duplex at rates between 1 and 2 kb per minute[19]. Although translocating at such high speeds, the parental strand is copied with high fidelity making an error only once about every 10^4 - 10^5 nucleotides polymerized[20]. Recently, a minimal set of proteins has been defined to reconstitute the *S. cerevisiae* replisome using only 16 purified replication factors, which has given us unprecedented insights concerning eukaryotic DNA replication[17]. Several other advantages, such as a defined origin of replication, have made studying DNA replication in the budding yeast *S. cerevisiae* particularly productive. In this section some of the crucial processes during DNA replication important to understand the goal and some of the methods used during this study will be described.

T.1.1 Loading of the minichromosome maintenance complexes

Eukaryotic DNA replication is initialized at many origins of replication across the genome. To prevent origins from firing multiple times, which results in erroneous replication, this process is tightly regulated ensuring the genome is duplicated only once per cell cycle.

During G1-phase the preparations for replication occur and pre-replicative complexes (pre-RC) consisting of ORC, Cdc6, Cdt1, and MCM proteins are formed at origins of replication. In *S. cerevisiae,* origins of replication are defined by the ARS1 consensus sequence. Recently, it was shown that ORC rapidly diffuses across the parental DNA duplex and is halted by origin recognition[21]. The ORC complex consists of six protein subunits (Orc1-6) and its main function is to facilitate and recruit other factors involved in the establishment of pre-RCs (**Fig. 1**)[22].

MCMs, precursors of the functional helicase complex of the replisome, are loaded head-to-head (N-termini facing each other) onto origins of replication, forming a DH which encircles the parental DNA duplex. The opposite orientation of the loaded MCMs anticipates the establishment of bidirectional replication forks. The loading process is carried out by the combined action of ORC, Cdt1, and Cdc6, which together with the MCMs form the ORC-Cdc6-Cdt1-Mcm (OCCM) complex, during a process commonly known as origin licensing (**Fig. 1**)[6].

The MCM complex is a heterohexameric ring shaped helicase which has a positively charged central channel that accommodates the negatively charged DNA duplex. The complex contains six AAA+ ATPase domains which are located at the interface between neighboring subunits. Loading of the MCMs is accompanied by various conformational changes. The helicase ring needs to be opened to be able to encircle the parental DNA duplex. Mcm2 and Mcm5, two subunits of the MCM complex, function as a gate which can open and close making it possible for the complex to encircle the parental DNA duplex. Cdt1 enhances the preference of the complex to be in the open state through the formation of a Cdt1-MCM heptamer that is recruited by ORC-Cdc6 during the formation of the OCCM complex[23].

ATP binding and hydrolysis play an important role in helicase loading, since ATP binding by ORC and Cdc6 is required for recruitment of the Cdt1-MCM heptamers to the origins of replication. To complete the loading process ATP hydrolysis is required after which the MCM ring closes and Cdt1 is released[24].

Subsequent steps leading to initiation of DNA replication during S phase have been shown to be independent of ORC and Cdc6 in *Xenopus laevis* egg extracts, supporting the claim that the main function of these factors is loading of the MCMs during replication initiation[25].

Although pre-RCs are established at all origin of replication sites, only few are activated to become functional replisomes. This is commonly referred to as the MCM-paradox[26]. It is thought that the purpose of excess MCMs on 'dormant origins of replication' is to act as back-up when replisomes are stalled in response to replicative stress[9].

T.1.2 Rearrangement from the loaded double hexamers into active replisomes

During S-phase of the cell cycle, the previously loaded pre-RCs are further assembled into active replisomes. These active replisomes are centrally organized by the Cdc45-Mcm2-7-GINS (CMG) helicase, which consists of 11 protein subunits. The activation involves the recruitment of Cell division cycle 45 (Cdc45) and the heterotetrametric GINS (go-ichi-ni-san, composed of Sld5, Psf1, Psf2, and Psf3) complex[27]. Besides Cdc45 and GINS, assembly and activation of the CMG complex requires an additional set of proteins known as firing factors.

The MCMs remain stably bound to ADP after loading but the firing factors trigger its release and subsequent binding of ATP promotes stable CMG formation[8]. To become fully active, the MCM ring must re-open and one of the two parental duplex strands must be removed from the central channel of the complex, whilst the other is retained. Finally, after activation by the firing factors, the CMG helicases translocate in an N-terminus first direction passing each other within the origin[8].

Instead of having to load inactive MCMs which need to be assembled into active CMG complexes by an ensemble of loading and firing factors, purified CMG from the yeast system allows the direct loading of CMG complexes onto DNA. This is very useful for single-molecule studies, however, to load the purified CMG complexes a ssDNA 3' overhang is required onto which CMG can bind and start translocating in the 3' to 5' direction, fuelled by the hydrolysis of ATP. To this end, two fluorescently labelled 3' polyT tails are incorporated into the construct synthesized in this study, which simultaneously function as a nanoruler for calibrating length measurements relative to the center of the construct of in bulk loaded proteins.

T.1.3 Replication fork convergence and bypass

During replication termination (**Fig. 4**), two converging CMG complexes normally bypass each other, after which the replication process is almost complete[28]. After the encounter, the Mcm7 subunit of the CMG complexes are polyubiquitylated and CMG is unloaded from the chromatin[12,13]**.**

It has been shown that CMG unloading only occurs after the gap between the leading strand of one replicative fork and the lagging strand of the other replicative fork has been filled and ligated, a process commonly referred to as gap filling. Recently is was shown that the polyubiquitination of CMG is impaired by its association with a replication fork, dependent upon the presence of the excluded DNA strand[29]. After two converging CMG complexes have passed each other, they translocate along the leading strand template until they encounter the downstream Okazaki fragment, whereupon the CMG complexes pass onto dsDNA, which is thought to allow the polyubiquitylation and subsequent unloading of the CMG complexes[16,29].

However, the unloading process does not always happen smoothly. Inactive DHs may obstruct CMG convergence and therefore inhibit successful replication termination (**Fig. 2**). CMG is able to push the inactive DHs forwards whilst translocating[8]. Obstruction of converging CMGs by DHs has not been shown directly. However, the DNA construct synthesized in this study will allow us to study this directly on the single-molecule level.

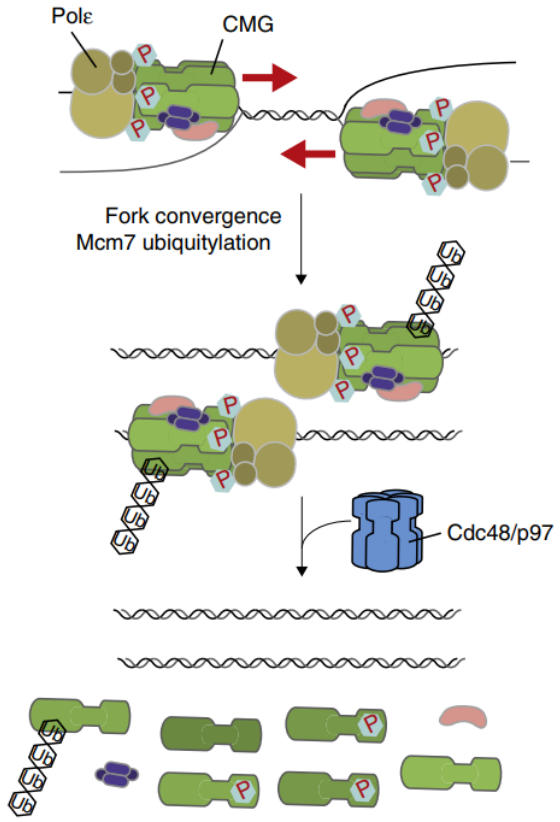


***Figure 4**. Schematic overview of the CMG unloading process, during replication termination. First the CMG complexes bypass each other, subsequently the Mcm7 subunits of the CMG complexes are polyubiquitinylated, and CMG is unloaded from the DNA by Cdc48/p97. Figure taken from Deegan & Diffley, 2016.*

Recently, it was shown that auxiliary DNA helicases, Pif1 and Rrm3, were found to stimulate fork convergence during DNA replication termination[14]. The accessory helicases jointly help the replisome to bypass obstructions at many sites across the genome, such as tRNA promoters where non-nucleosomal proteins are found to tightly bind the DNA. Furthermore, Hill et al. (2020) showed that in the presence of in excess loaded MCMs, successful replication termination is inhibited because of the inactive MCMs. However, when Pif1 was added as well, almost all obstructing MCMs were removed and the termination process was restored, showing Pif1 plays a crucial role in aiding successful replication termination. The removal of DHs by Pif1 has not been shown directly, but the construct synthesized in this study in combination with the bacterial overexpression plasmid containing a single-solvent-exposed-cysteine mutant of Pif1 allows to study its interaction with the obstructed and converging CMG helicases during termination.

T.2 Worm-like chain models

To biophysically characterize the DNA construct that we synthesise, force-extension data obtained from an optical trapping experiment is fitted using the inextensible and extensible worm-like chain models. At low forces, between 0.01 and 10 pN, bends from the DNA are removed and the DNA duplex can be modelled as an entropic spring[30]. This force-extension regime should be well described by the inextensible worm-like chain (WLC) model, which takes into account the energetic cost of bending.

The polymer flexibility is characterised by the persistence length, the distance over which two segments of the chain remain directionally correlated. It has been shown that a high salt concentration influences the persistence length. The positively charged ions will shield the negative charges from the DNA backbone, causing it to experience less electrostatic repulsion and thereby lowering the energetic penalty for bending, and causing the persistence length to decrease[31,32].

Around 10 pN, when the DNA is stretched slightly beyond its B-form contour length, increasingly higher and higher forces are required to extend the DNA. In this regime the force-extension data should be well described by the extensible worm-like chain (xWLC) model which includes an elastic stretch modulus which accounts for the elastic response resulting from stretching the DNA beyond its B-form contour length, frequently referred to as overstretching. The origin of the overstretching transition has been attributed to a secondary structure transition of the DNA from B-form to a stretched S-form. Finally, when the DNA is stretched even further the DNA suddenly yields in a highly cooperative fashion and overstretches about 1.7 times, the observed plateau in the force-extension curve signals the cooperative overstretching transition and occurs at about 60 to 70 pN[31].

In this study **equation 1** and **equation 2** taken from Bouchiat et al. (1999) are used to fit the force-extension data to the WLC and xWLC model, respectively. From the WLC model fit a persistence length (L_p) and a contour length (L_0) can be obtained. The xWLC yields an additional elastic stretch modulus (K_0). Matlab scripts which were already present in the laboratory were converted and adapted to Python 3.8 to help the laboratory streamline the data analysis.

$$\boldsymbol{equation\ 1.} \quad \frac{FL_p}{k_bT} = \frac{1}{4}\left(1-\frac{x}{L_0}\right)^{-2} - \frac{1}{4} + \frac{x}{L_0} + \sum_{i-2}^{i-7} \alpha_i \left(\frac{x}{L_0}\right)^i$$

$$where\ \alpha_2 = -0.5164228, \alpha_3 = -2.737418, \alpha_4 = 16.07497, \alpha_5 = -38.87607, \alpha_6 = 39.49944, \alpha_7 = -14.17718$$

$$\boldsymbol{equation\ 2.} \quad \frac{FL_p}{k_bT} = \frac{1}{4}(1-l)^{-2} - \frac{1}{4} + l + \sum_{i-2}^{i-7} \alpha_i (l)^i$$

$$where\ l = \frac{x}{L_0} - \frac{F}{K_0}\ and\ coefficients\ \alpha_i\ same\ as\ above$$

T.4 Optical tweezers general principle

One of the main instruments used in this study is Lumicks C-trap® G2 which allows simultaneous manipulation and visualization of single-molecule interactions in real time by combining high resolution optical tweezers, fluorescence microscopy, and an advanced microfluidic system. The optical tweezers technique is based on the momentum of photons which is very small, but when bundled in rays, large enough to exert a force on micron-sized particles (**Fig. 5a**). These forces can be used to optically trap microscopic particles such as polystyrene beads, which in turn can exert forces on even smaller particles such as DNA (**Fig. 5b and c**). In this manner, optical tweezers can be used to manipulate particles on the molecular scale[33].

The forces exerted by the trapping lasers on the microspheres can be modelled by a Hookean spring system, with a characteristic trap stiffness which depends on the intensity of the trapping lasers. The Gaussian beam profile created by the trapping laser causes a force gradient. Any deviation from the middle of the laser beam causes an unequal force, which pushes the particle back into position. By

calibrating the setup, the spring constant can be determined and applied forces can then be calculated from the bead displacement.

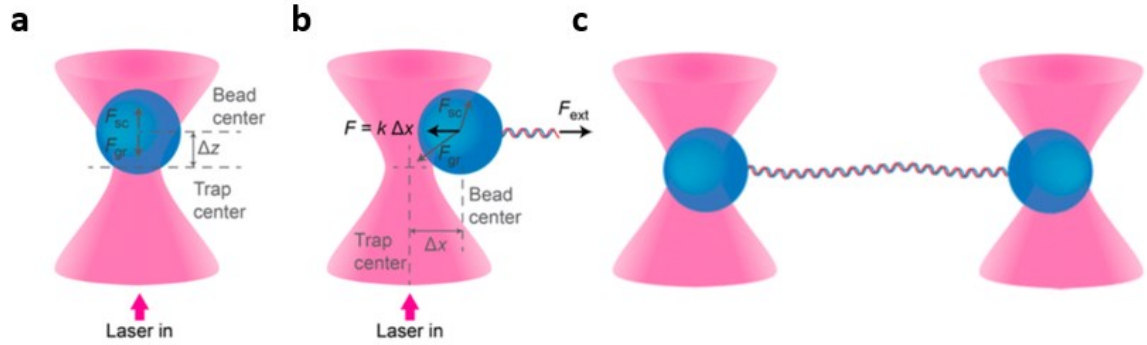


***Figure 5**. (**a**) An optically trapped microsphere. The scattering force acts in the direction of light propagation while the gradient force acts in the opposite direction along the electric field gradient, in this illustration both forces cancel each other resulting in a stable optically trapped microsphere. (**b**) An external force exerted on the microsphere causing a lateral displacement of Δx causes an opposing and restoring force which can be described by Hooke's law as F = kΔx, where k represents the trap stiffness. (**c**) Cartoon showing a DNA molecule trapped by two microspheres. Figure adapted from Heller et al., 2014.*

Materials and methods

M.1 DNA construct synthesis and purification

The synthesis of the desired construct can be divided into five steps (**Fig. 6**): plasmid linearization, biotinylation, nicking, fork annealing, and fork ligation. Here, the step-by-step protocol for synthesis of the construct will be described.

M.1.1 Step1; Plasmid linearization. To linearize the plasmids, 10 µg of plasmid pDRM2 were digested in 1X of 10X Cut Smart Buffer (NEB Cat. No. B7204S) at 37 °C overnight with 50 units of XhoI (NEB Cat. No. R0146S) in 50 µl final volume. Following the digestion, the restriction enzyme was heat-inactivated at 65 °C for 35 min.

M.1.2 Step 2; Biotinylation. After plasmid linearization the sticky ends are blunted using biotinylated deoxyribonucleotides by 15 units of Klenow Fragment (3' to 5' exo-) (NEB Cat. No. M0212L) in 10X NEBuffer™ 2 (Cat. No. B7002S) at 37 °C for 1 h using 33 nM Deoxyguanosine triphosphate, 33 nM biotin bound Deoxyadenosine triphosphate, 33 nM biotin bound Deoxycytidine triphosphate, and 33 nM biotin bound Deoxyuridine triphosphate in 91 µl final volume. Following incubation, the reaction is terminated by addition of 10 mM EDTA, subsequently Klenow fragment is heat-inactivated at 37 °C for 20 min. Finally, the mixture is flown through an illustra™ MicroSpin™ S400 spin column to remove any free unincorporated nucleotides according to the manufacturers' protocol but centrifuging at 0.8 RCF for 1 minute during column preparation and 2 min during elution instead.

M.1.3 Step 3; Nicking. Nicking is achieved by adding 50 units of Nb.BbvCI (NEB Cat. No. R0631L) in 10X NEBuffer™ 2 (Cat. No. B7002S) to the reaction mixture obtained after biotinylation in a final volume

of 105 µl. Depending on which protocol was followed this reaction is incubated for either 3 or 5 h at 37 °C. Following nicking, the enzyme is heat-inactivated for 20 min at 80 °C.

M.1.4 Step 4; Fork insertion. To achieve fork annealing the fluorescent fork oligonucleotides (either DRM 158 or DRM 159, **see appendix**) are added in 100x molar excess to the reaction mixture obtained after nicking. To remove the small fragments, which resulted from nicking, the mixture is heated to 80 °C for 10 min. The forks are then annealed to the complementary single stranded regions, by cooling the reaction down to 20 °C at a rate of 1 °C/min in a thermocycler (Bio-Rad # 1851148). To prevent photobleaching of the fluorophores this reaction is covered in tin foil.

M.1.5 Step 5; Ligation. The final step is to ligate the 5' end of the newly inserted fork oligonucleotides to the 3' end of the construct. To achieve this, 1600 units of T4 ligase (NEB Cat. No. M0202S) is incubated at 16 °C overnight with 1.54 mM ATP in 10X NEBuffer™ 2 (Cat. No. B7002S) and the reaction mixture obtained after fork insertion in a final volume of 130 µl. Subsequently, T4 ligase is heat-inactivated at 65 °C for 10 min.

M.1.6 Construct purification. To purify the construct, the solution obtained from construct synthesis is run on a 0.8% agarose gel for 1.5-2 h at 100 V in TAE using SybrSafe™ to stain the DNA. Following electrophoresis, the 10 kb band corresponding to the construct is excised and purified using the Promega's Wizard® Genomic DNA Purification Kit according to the manufacturers' protocol. The purified constructs are then stored at 4 °C.

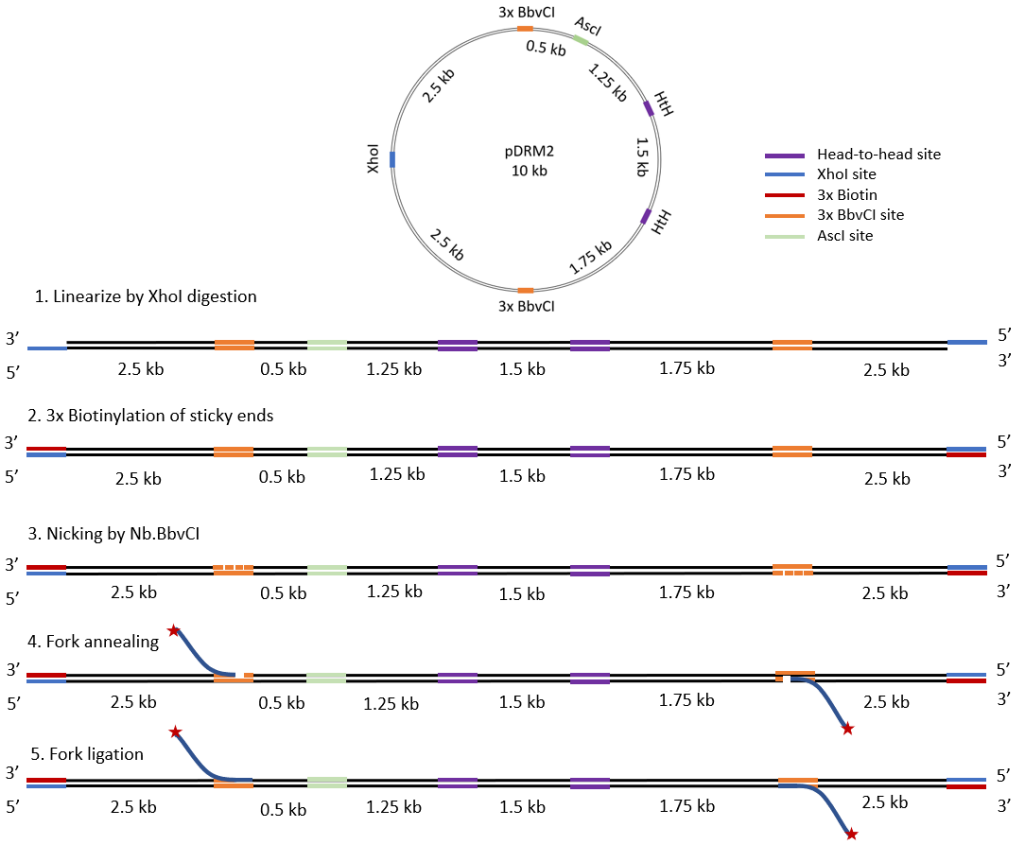


***Figure 6**. The five different steps in the synthesis process of the DNA construct. (**1**) The first step is to linearize the construct by digesting the plasmid using the XhoI restriction enzyme. (**2**) After digestion, the resulting sticky ends are biotinylated by Klenow fragment. (**3**) Following biotinylation, the construct is nicked at the fork insertion sites by Nb.BbvCI. (**4**) During step 4, the small fragments resulting from nicking are removed, and the fluorescent fork oligonucleotides are annealed to the single stranded*

*regions. (**5**) The last step in the synthesis process is the ligation of the newly inserted fork oligonucleotides to the rest of the construct by T4 ligase.*

M.2 Bulk loading reactions

M.2.1 gRNA folding & dCas9 loading reactions in bulk. gRNA folding was achieved as following: 2 µM crRNA DRM 1 (**see appendix**) and 2 µM tracrRNA were mixed in 1X IDT Nuclease-Free Duplex Buffer in 20 µl final volume, placed in a beaker with water at 100 °C and allowed to slowly cool down to room temperature overnight. The dCas9 loading reactions in bulk were carried out as follows: 25 nM of either dCas9-Halo-AF488 or dCas9-Halo-JF646 was incubated with 250 nM gRNA (gRNA DRM 1 or gRNA DRM 2, **see appendix**) for 10 min at 30 °C and 1250 RPM mixing. Subsequently, the reaction mixture was incubated with 3.3 nM DNA substrate in 1X of 2X loading buffer (50 mM HEPES-KOH [pH7.6], 20 mM Magnesium acetate [MgAc], 200 mM potassium glutamate [KGlu], 0.04% NP40, 20% of 50% Glycerol, 2 mM Dithiothreitol [DTT], 0.2 mg/ml BSA, and 10 mM ATP-Mg) for 40 min at 30 °C.

M.2.2 MCM loading reactions in bulk. Loading reactions were carried out as follows: 10 nM ORC was incubated with 3.3 nM DNA substrate (with or without pre-loaded dCas9, depending on the experiment) for 5 min at 30 °C in 1X of 2X loading buffer (See dCas9 loading reactions) at 1250 RPM mixing. Subsequently, the mixture was incubated with 10 nM Cdc6 for 5 min at 30 °C and 1250 RPM mixing. Finally, the mixture was incubated with 20 nM [JF646-Halo-Mcm3]Mcm2-7-Cdt1 for 20 min at 30 °C and 1250 RPM mixing in a final volume of 20 µl.

M.3 C-trap sample preparation, cleaning protocol, and data acquisition

M.3.1 Sample preparation. The imaging buffer was obtained by mixing the following substances: 1X of 2X loading buffer (see dCas9 loading reactions), 2 mM 1,3,5,7 cyclooctatetraene, 2 mM 4-nitrobenzylalchohol, 2 mM Trolox, 2.5 mM Protocatechuic acid, and 10 mM Protocatechuate-3,4-dioxygenase in 1 ml final volume. Protocatechuate-3,4-dioxygenase together with Protocatechuic acid and Trolox function as an oxygen-scavenger system which prolongs the lifetime of the fluorophores. To obtain the bead sample 1.0% w/v of streptavidin-coated polystyrene beads with a diameter of 1.76 µm are sonicated for 20 min at room temperature in a water bath sonicator. To sterilize the buffers these are filtered using a 0.2 µm filter and subsequently degassed for 20 min.

M.3.2 C-trap cleaning protocol. To clean the C-trap, the contents of the syringes is removed using a pipet, and 1 ml of bleach is flown through for 45 min at 0.5 bar. The remaining bleach is then removed from the syringes using a pipet and 1 ml of MilliQ is flown through for 45 min at 0.5 bar. Afterwards, the remaining MilliQ is removed from the syringes using a pipet and 10 nM sodium thiosulfate in 1.5 ml final volume, is flown through for 45 min at 0.5 bar. Finally, the remaining sodium thiosulfate is removed from the syringes using a pipet and 1 ml of MilliQ is flown through at 0.5 bar for 45 min. After the final step, the syringes are left in MilliQ until usage. In case of a blockade (e.g. a bubble) the pressure is temporarily increased to 1 bar to remove it.

M.3.3 C-trap data acquisition. The microfluidic flow cell of the C-trap provides six laminar flow channels with a single exit channel. In this study, three of the six flow channels are used (**Fig. 7**). The first channel contains a solution with streptavidin-coated polystyrene beads (diameter of 1.76 µm). The second channel contains the synthesized DNA constructs with or without proteins bound to it, depending on the experiment. In the third channel, an imaging buffer is dispensed which contains the anti-photobleaching cocktail described above.

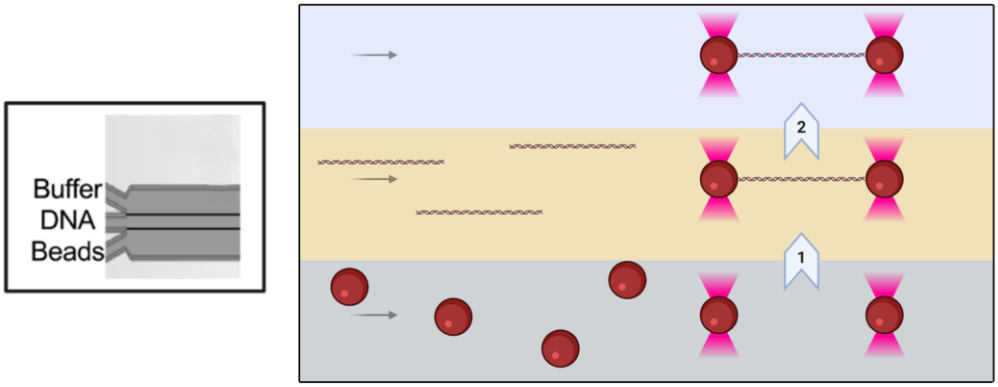


Figure 7. *Schematic representation of the microfluidic flow cell. (**Bottom**)In the first channel a solution containing the streptavidin-coated polystyrene beads (diameter of 1.76 µm) is dispensed. (**Middle**) In the second channel a solution containing the DNA constructs is dispensed. (**Top**) In the third channel an imaging buffer is dispensed. Figure courtesy of Sánchez H.*

To acquire data, a construct molecule was trapped as following: Streptavidin-coated polystyrene beads (diameter of 1.76 µm) are optically trapped in the first channel by the trapping lasers. The laser power of the lasers is adjusted to achieve a trap stiffness of 0.3 $pNnm^{-1}$. In the second channel, construct molecules are trapped, a sudden rise in the force-extension curve indicates presence of a trapped particle. Subsequently, in the third channel containing the imaging buffer, force-extension data and fluorescence scans of the trapped particle are acquired. All measurement were carried out under a 2 pN force-clamp, to image all DNAs under approximately the same extension. The field of view of 5.5 by 1 µm included half of both beads. The time it took to scan one pixel, one horizontal line of pixels, and the complete field of view are 0.05 ms, 9.6 ms, and 0.6 s, respectively. Lastly, the laser powers used for the blue, green, and red lasers of the confocal microscope are 0.42 µW, 0.714 µW, and 0.35 µW, respectively.

M.4 Generation of bacterial overexpression plasmid containing a single-solvent-exposed-cysteine mutant of Pif1. To generate a single-solvent-exposed-cysteine mutant of Pif1, a second mutation (C550S) was introduced into the pET28a-Pif1(40-859) C80S plasmid previously generated (Daniel Ramírez Montero, personal communication). This is achieved as following: 0.4 ng/µl of pET28a-6His-Pif1(40-859) C80S is mixed with 0.5 µM Primer for both the forward and reverse strands (DRM 162 and DRM 163, **see appendix**) in NEB's Q5® Hot Start High-Fidelity 2X Master Mix , in a final volume of 25 µl. The plasmids are then mutated, and PCR amplified by heating the sample to 98 °C for 30 s in the thermocycler. Subsequently, the sample is kept at 98 °C for 10 s cooled to 56 °C for 30 s, and heated to 72 °C for 5 min and 30 s. This cycle is repeated 29 times, after which the sample is kept at 72 °C for 2 min and finally at 12 °C until removed. After mutagensis the PCR products are circularized by ligation and the parental pET28a-Pif1(40-859) C80S plasmids are degraded by incubating 1 µl PCR product, in 1X of 2X KLD reactive buffer and 1X of 10X KLD enzyme mix (NEB Cat. No. M0554S) in a final volume of 10 µl for 10 min at room temperature.

M.5 Data analysis

M.5.1 Particle localization in 2D scans. The open-source single-particle tracking ImageJ plugin TrackMatev5.2.0[35] was used to identify and track the positions of particles bound to the constructs

over time during the single-molecule experiments. The plugin produces a set of coordinates (x, y, t) for identified regions of enhanced fluorescence. The positional and temporal data from the detected foci was then exported from TrackMatev5.2.0 to Python 3.8 (with analysis libraries numpy 1.19.0, scipy 1.5.0, scikit-image 0.17.2, scikit-learn 0.23.1, and pandas 1.05) as XML files.

M.5.2 Localization distributions relative to the DNA center. To determine the positional distribution of particles with respect to the center of the DNA construct, the position of the center had to be determined. From the ensemble of raw localization distributions of all scans we consistently obtained four distinct peaks. The field of view contained half of both beads, therefore the two outer peaks (**Fig. 8**) were identified as false positives at the bead edges, which resulted from fluorescent proteins adhered to the beads. Furthermore, from the green color channel (Cy3) the inner two peaks were identified as peaks corresponding to the positions of the fluorescent fork oligonucleotides (**Fig. 8**). To find their mean position a Gaussian mixture model was fitted.

From the biochemical experiments we know the fork oligonucleotides are symmetrically located around the center, therefore the center of the mean fork positions was identified as the mean center of all optically trapped constructs. All distances are then expressed relative to this center. All molecules are trapped and scanned under similar conditions and are symmetrically located around the center of the construct, therefore we think averaging over the ensemble of trapped particle positions is justified.

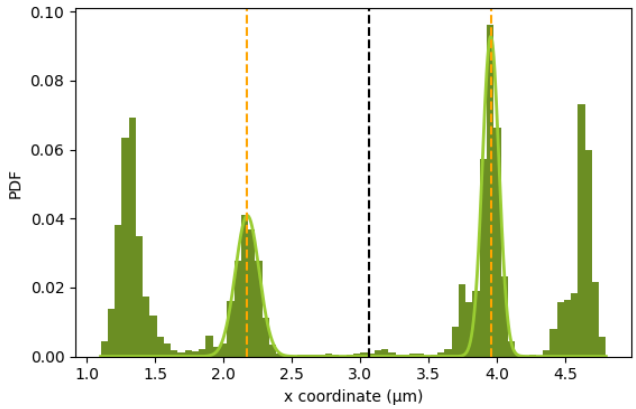


***Figure 8**. Exemplary histogram, showing how the center is determined from the ensemble of raw localization distributions of all scans from the green color channel (Cy3) for one of the experimental data sets. The two outer peaks are identified as false positives at the bead edges, whereas the middle two peaks correspond to the fork oligonucleotides. The mean positions of the fork oligonucleotides is found by fitting a Gaussian mixture model to the inner two peaks and are demarcated by the vertical orange dashed lines. The mean center is identified as the mean of fitted fork positions and is demarcated by the vertical black dashed line.*

Furthermore, using the distance between the positions of the fork oligonucleotides a micron-to-base-pair coordinate conversion factor is determined. From the biochemical experiments we know the distance between both fork insertion sites is 5 kb. This conversion factor was subsequently used to express the positions of proteins relative to center of the construct in kb and compare the values with the expectations.

M.5.3 Survival probability analysis. To calculate the mean lifetime of the fluorophores the total amount of detected foci per frame over the ensemble of all scans was fitted to the following equation:

$$N(t) = N_0 e^{\frac{-t}{\tau}}, where\ N_0 = N(0), and\ \tau = mean\ lifetime$$

which was done in Python 3.8 (with analysis libraries numpy 1.19.0 and scipy 1.5.0).

Results and discussion

R.1 Plasmid assembly and characterization

R.1.1 Plasmid assembly, transformation, and initial screen of bacterial clones

To synthesize the construct for the single molecule experiments, the desired plasmid containing the required restriction sites at the desired locations first had to be assembled and amplified. This is because it was not possible to engineer pre-existing plasmids available in the laboratory into the desired plasmid in a timely fashion.

The plasmid was assembled by Gibson assembly in a NEBuilder reaction (NEB Q5® Hot Start High-Fidelity DNA Polymerase kit) using four custom designed gBlocks™ of 2.5 kb with overlapping ends. The overlapping ends have a melting temperature of $T_m \approx 60$ °C and the gBlocks™ are fabricated by IDT. To assemble the gBlocks™ into a plasmid (pDRM2) (Daniel Ramírez Montero, personal communication), 2 fM of each gBlock™ was incubated with 1X of 2X HiFi MasterMix in a final volume of 20 µl at 50 °C for 2 h. Subsequently, the plasmids were transformed into NEB® 5-alpha Competent *E. coli* (High Efficiency) cells (NEB Cat. No. 02987) and plated on selection media.

From the transformation, four colonies were obtained, and these were subjected to an initial restriction enzyme digestion screen to assess whether the gBlocks ™ were correctly assembled. The DNA from each colony was purified using the Promega PureYield™ Plasmid Miniprep System, according to the manufacturers' protocol, and was subsequently digested in three different test reactions. These reactions included either the individual or combined cleavage by the restriction enzymes XhoI and BbvCI. In each reaction, 175 ng of DNA was digested at 37 °C for 1 h and 36 min using either 10 units of XhoI or 1 unit of BbvCI for the single digestions, or of both for the combined reactions. Not all different combinations of test digestions were conducted, because little DNA was obtained from the DNA purification using the miniprep system, and these digestions are only meant as an initial screening.

Together with an undigested sample, the digested plasmid was run on a 0.8% agarose gel in TAE using ethidium bromide to stain the DNA. After this initial screen, two clones whose band patterns matched the expected pattern were selected to amplify and purify the plasmids in a larger scale for further characterization and to synthesize the DNA construct for the single-molecule experiments.

R.1.1.1 Expected band patterns from initial screen of bacterial colonies

Cleavage of the desired plasmids by XhoI is expected to result in linearized 10 kb fragments (**Fig. 9**). Furthermore, digestion by BbvCI is expected to result in one band of approximately 5 kb. The combined action of XhoI and BbvCI on the desired plasmid is expected to yield two fragments of 2.5 kb and a fragment of 5 kb. Digestion by BbvCI should in theory result in four additional fragments of approximately 15 bp, because both BbvCI clusters contain three BbvCI cut sites which are separated by approximately 15 bp. Ethidium bromide intercalates between the nitrogenous bases of DNA, therefore the intensity is proportional to the length of the DNA. Therefore, the short fragments are not expected to be visible on the gel, because their intensity may be too low compared to the longer bands to be able to detect. Furthermore, the short fragments may run off the gel, and will diffuse at an enhanced rate resulting in unsharp bands. The expected band patterns corresponding to these digestion reactions are shown in the simulated gel (**Fig. 10b**) and **Table 1**.

Restriction enzymes	Expected fragments after full digestion	Possible additional fragments due to partial digestion
XhoI	10 kb linear	10 kb nicked, 10 kb supercoiled
BbvCI	5 kb	10 kb all conformations
XhoI + BbvCI	2.5 kb, 5 kb	7.5 kb, 10 kb all conformations

***Table 1**. Expected fragments after a full or partial digestion of the desired plasmid using different combinations of restriction enzymes, 10 kb all conformations includes: open linear, circular relaxed, and supercoiled.*

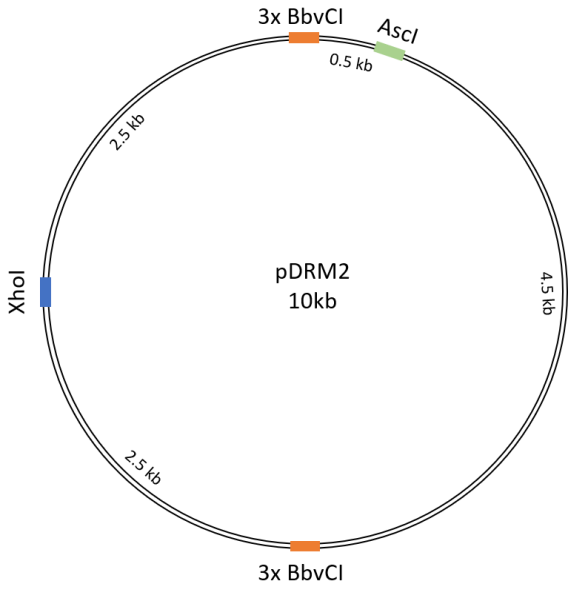


***Figure 9**. Map of the desired 10 kb plasmid (pDRM2) showing only the location of the different restriction sites. The recognition sequence for Nb.BbvCI and BbvCI is the same, therefore the BbvCI sites can also be nicked by Nb.BbvCI.*

The undigested plasmid serves as negative control and can be present in three different conformations: supercoiled, open linear, or circular relaxed (commonly known as nicked). The supercoiled state is the most compact of these three conformations and hence experiences the least amount of drag when moving through the gel matrix. Therefore, it is expected to run lower than the plasmids in any of the other conformations. When the plasmid is nicked at one of the two strands, circularity will remain, but the breakage will permit rotation around the phosphodiester bonds, causing the supercoils to be relaxed. The nicked conformation is therefore less compact than the super coiled conformation, hence it is expected to run higher on the gel. Lastly, the open linear conformation is also expected to run higher than the supercoiled conformation[36]. In the undigested sample, it is expected to see all three different conformations, because during plasmid propagation and purification, some of the plasmid becomes cleaved or nicked by endogenous nucleases and/or shearing forces involved in the handling.

R.1.1.2 Reactions were not run to completion, faint smears are likely due to degradation by DNases, and nicked conformation runs higher than open linear conformation of the plasmid

The most intense band patterns of colonies C1, C3, and C4 (**Fig. 10a**) correspond to the expected result. However, there are some additional unexpected bands present which can be accounted for by partial or no cleavage. From the diagnostic gel we see the reactions were not run to completion. This can be seen by looking for example at lane 5 were there are undigested bands present corresponding to the open linear conformation of the plasmid, whereas complete digestion would result only in fragments of 5 kb. Determining the exact time to run all reactions to completion is difficult, because the activities

of the enzymes depend on multiple factors (e.g. temperature, concentration, buffer, etc.). Moreover, it is not needed to run the reactions till completion to determine whether the clones seem to have the desired plasmid. Therefore, there are still some reaction intermediates present in which the plasmid is partially cleaved or not at all, hence the band patterns do not correspond fully to those of the simulated gel (**Fig. 10b**), which assumes complete digestion.

Some lanes of the diagnostic gel (**Fig. 10a**) do not contain any bands, only a faint smear at the bottom of the gel. This was not expected and is probably caused by the introduction of DNases during the setup of the reactions. The faint smear at the bottom of these lanes are therefore likely degradation products.

Furthermore, the nicked plasmid conformation runs higher than the open linear conformation, because the band corresponding to the open linear conformation should run as high as the most intense band from the samples digested by XhoI. By looking at **lane 2** (**Fig. 10a**) it can be seen there is a band running higher than the band corresponding to the open linear conformation, which is likely the nicked plasmid. Below the band patterns obtained from each clone will be described in order and in detail.

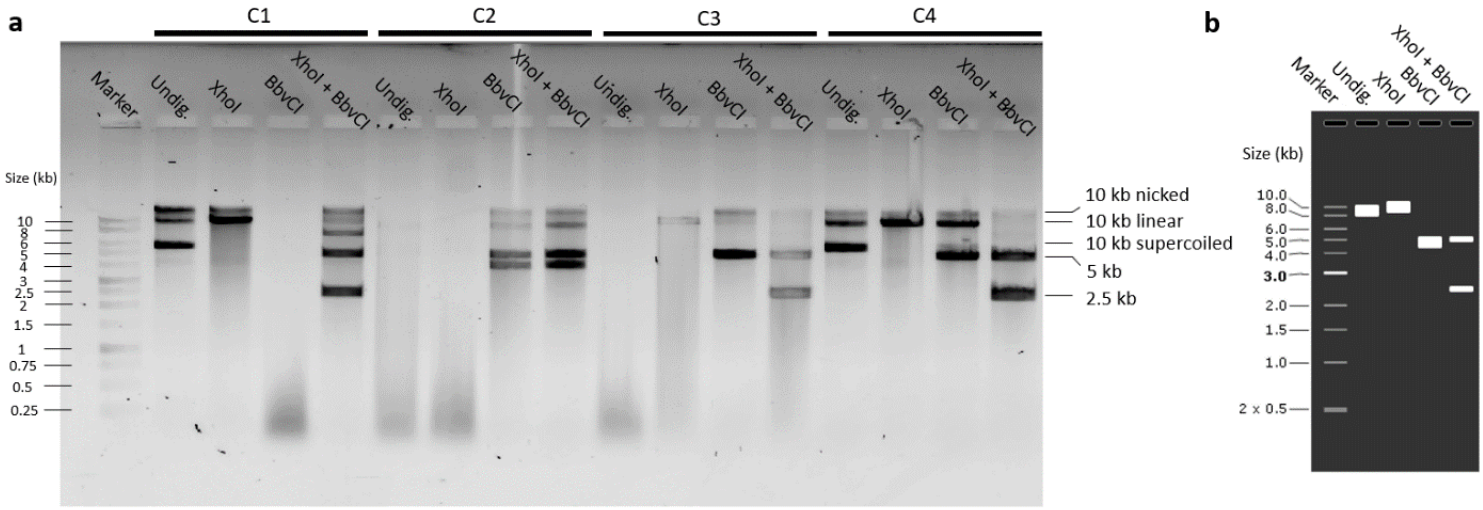


***Figure 10**. Diagnostic gel obtained after initial screening of the obtained colonies. (**a**) Per clone, three different digestions were conducted using the combined or individual action of the XhoI and BbvCI restriction endonucleases. (**b**) Gel simulation of the expected band patterns, assuming the digestion reactions run till completion. The simulation was made using SnapGene®5.2.4.*

R.1.1.3 Clone 1 seems to have the desired plasmid

The XhoI digestion of C1 (**Fig. 10a, lane 3**) seems to have an additional band which runs slightly higher than the band corresponding to the linearized 10 kb fragment. This band is caused by the circular relaxed conformation of the plasmid, because it is also present in the negative control where it corresponds to the circular relaxed band pattern. In a complete digestion reaction, a nicked version of the uncut plasmid would not be present, because all the plasmids would be cleaved. Furthermore, the band pattern of the combined digestion of C1 by XhoI and BbvCI (**Fig. 10a, lane 5**) shows three additional bands which are not present on the simulated gel. One band having a length of approximately 7.5 kb, which can be explained by partial cleavage of the plasmid by BbvCI in one instead of both BbvCI restriction site clusters, combined with cleavage of the XhoI site. The other additional bands in this lane correspond to the open linear and circular relaxed conformation of the

plasmid. Although the data of the digestion with only the BbvCI restriction enzyme is missing, it is still possible to draw the conclusion that the BbvCI sites are present at the desired locations by looking at the results of the combined digestion of BbvCI and XhoI. To conclude, all the band patterns of the digested plasmid DNA obtained from C1 can be explained by digestions of the desired plasmid. Therefore, C1 likely contains a plasmid of the expected size with the correct sequence.

R.1.1.4 Clone 2 seems to have a plasmid with an incorrect order of gBlocks

The results after cleavage by BbvCI of plasmids obtained from C2 are unexpected (**Fig. 10a, lane 8**) There seems to be a band of 7.5 kb and a band of about 4-4.5 kb. These bands cannot be explained due to partial cleavage. Furthermore, the band patterns obtained after digestion by XhoI combined with BbvCI (**Fig. 10a, lane 9**), were not expected either. It seems to be an exact copy of the lane before it. In this case the 7.5 kb fragment can be explained by partial cleavage. However, the band at 4-4.5 kb still cannot be explained by any possible combination of cleavage by these two restriction enzymes on the desired plasmid. Together, these data suggest that a different undesired combination or order of gBlocks was transformed into colony 2. The degradation of DNA in half of the samples of C2, together with the fact that the lanes of the digestion with BbvCI and the digestion of XhoI combined with BbvCI seem to be identical, suggest that mistakes have been made during the setup of the reactions for this colony. To conclude, the results obtained from C2 do not correspond to the expected result and cannot be explained by cleavage of the desired plasmid. Therefore, C2 was discarded.

R.1.1.5 Clones 3 and 4 seem to have the desired plasmid

The only differences with the simulated gel and the results obtained from C3, are the two faint bands corresponding to the linearized and the nicked version of the 10 kb plasmid obtained after digestion with BbvCI (**Fig. 10a, lane 12**). Although the lane with the negative control is missing, it can still be deduced from the other lanes that the plasmid is approximately 10 kb in length. In the band pattern obtained from digestion by BbvCI of C4 (**Fig. 10a, lane 15**), a band of about 10 kb can be seen, which is probably the linearized plasmid. Furthermore, the faint bands running higher than 10 kb are likely due to the circular relaxed conformation of the plasmid. Together, these data suggest that C3 and C4 contain plasmids of the expected size with the correct sequence.

R.1.2 Full biochemical characterization of clones 3 and 4

After analysis of the results obtained from the initial screening, colonies C3 and C4 were chosen to be studied in more detail, because they seemed the most promising. To assess the correct assembly of the plasmid, a more detailed characterization of these clones was conducted. The DNA from 200 ml cultures of these clones was purified using Marchery-Nagel's NucleoBond Xtra Midi kit, according to the manufacturers' protocol. To screen the colonies, eight different restriction digestions were conducted per clone. These reactions included either the individual or combined cleavage of the plasmid by XhoI, BbvCI, and AscI. Besides the digestions with XhoI in combination with BbvCI and their expected outcomes as described in the previous section, five additional digestions were conducted per colony to get a more detailed analysis of the clones. The undigested plasmid was used as a negative control and a digestion with Nb.BbvCI, a mutated version of the BbvCI enzyme which nicks DNA, was conducted to confirm nicking activity at the desired sites. In every reaction, 400 ng of DNA was digested at 37 °C for 2 h using 10 units of XhoI, 1 unit of BbvCI, or 5 units of AscI for the single

digestions or of all three of the enzymes for the combined reactions. The reaction products were run on a 0.8% agarose gel in TAE buffer stained with ethidium bromide.

R.1.2.1 Expected band patterns from full biochemical characterization of clones 3 and 4

The digestions with only AscI or XhoI are expected to linearize the plasmid, because only one corresponding restriction site is present on the desired plasmid. Therefore, the reaction product is expected to correspond with the linearized 10 kb band on the gel. As stated previously the Nb.BbvCI is a mutated version of the BbvCI restriction enzyme, it does not cleave the DNA, but instead nicks it on one of the two strands. Because Nb.BbvCI nicks the plasmid, the expected band corresponds with the circular relaxed conformation of the plasmid. The combined action of XhoI and AscI is expected to produce two fragments: a band of 3 kb and one of 7 kb in length. The reactions in which BbvCI is combined with AscI are expected to produce fragments of 0.5 kb, 4.5 kb and 5 kb. Lastly, the digestions in which XhoI is combined with AscI and BbvCI are expected to produce fragments of 0.5 kb, 2.5 kb and 4.5 kb. The expected fragments after full digestion in each reaction can be seen in **Table 2** and in the simulated gel of **Figure 11b**.

Restriction enzymes	Expected fragments after full digestion	Possible additional fragments due to partial digestion
XhoI	10 kb linear	10 kb nicked, 10 kb supercoiled
AscI	10 kb linear	10 kb nicked, 10 kb supercoiled
BbvCI	5 kb	10 kb all conformations
Nb.BbvCI	10 kb nicked	10 kb supercoiled, 10 kb linear
XhoI + AscI	3 kb, 7 kb	10 kb all conformations
XhoI + BbvCI	2.5 kb, 5 kb	7.5 kb, 10 kb all conformations
BbvCI + AscI	0.5 kb, 4.5 kb, 5 kb	5.5 kb, 9.5 kb, 10 kb all conformations
XhoI + BbvCI + AscI	0.5 kb, 2.5 kb, 4.5 kb	3 kb, 5 kb, 5.5 kb, 7 kb, 7.5 kb, 9.5 kb, 10 kb all conformations

Table 2. *Expected fragments after full or partial digestion of the desired plasmid using different combinations of restriction enzymes, 10 kb all conformations includes: open linear, circular relaxed, and supercoiled.*

R.1.2.2 Clone 3 likely contains the correct plasmid

The BbvCI digestion of the purified DNA from C3 (**Fig. 11a, lane 5**) contains an unexpected band of about 6 kb. This band is likely due to the supercoiled version of the plasmid, because it is also present in the negative control (**Fig. 11a, lane 2**), where it corresponds to the supercoiled conformation. In **lane 6** there is an additional unexpected faint band present corresponding to fragments of 5 kb. Synthetic restriction endonucleases such as Nb.BbvCI sometimes cleave sequences similar to their defined recognition sequence when the reaction is run under sub-optimal conditions, which we confirmed with the manufacturer. This is commonly referred to as 'star activity'. Therefore, the unexpected 5 kb fragments in lane 6 are likely due to star activity of the nicking enzyme. There are

many factors which influence these conditions and therefore it is difficult to get rid of star activity entirely. Furthermore, **lane 8** (**Fig. 11a**) contains an additional band of 7.5 kb and a band corresponding to the open linear conformation of the plasmid. The 7.5 kb band may be caused by cleavage in only one of the two BbvCI restriction clusters combined with cleavage at the XhoI site, which would result in fragments of 2.5 kb and 7.5 kb. In **lane 9** (**Fig. 11a**)**,** an additional band of 5.5 kb is present which can be explained by partial cleavage of the plasmid. Cleavage at the AscI site and a site in the BbvCI cluster nearest to the AscI site would result in two fragments, one of 4.5 kb and one of 5.5 kb. Bands corresponding to both fragments are present in this lane. Taken together, all band patterns are as expected or can be accounted for, showing that clone 3 likely contains the desired plasmid.

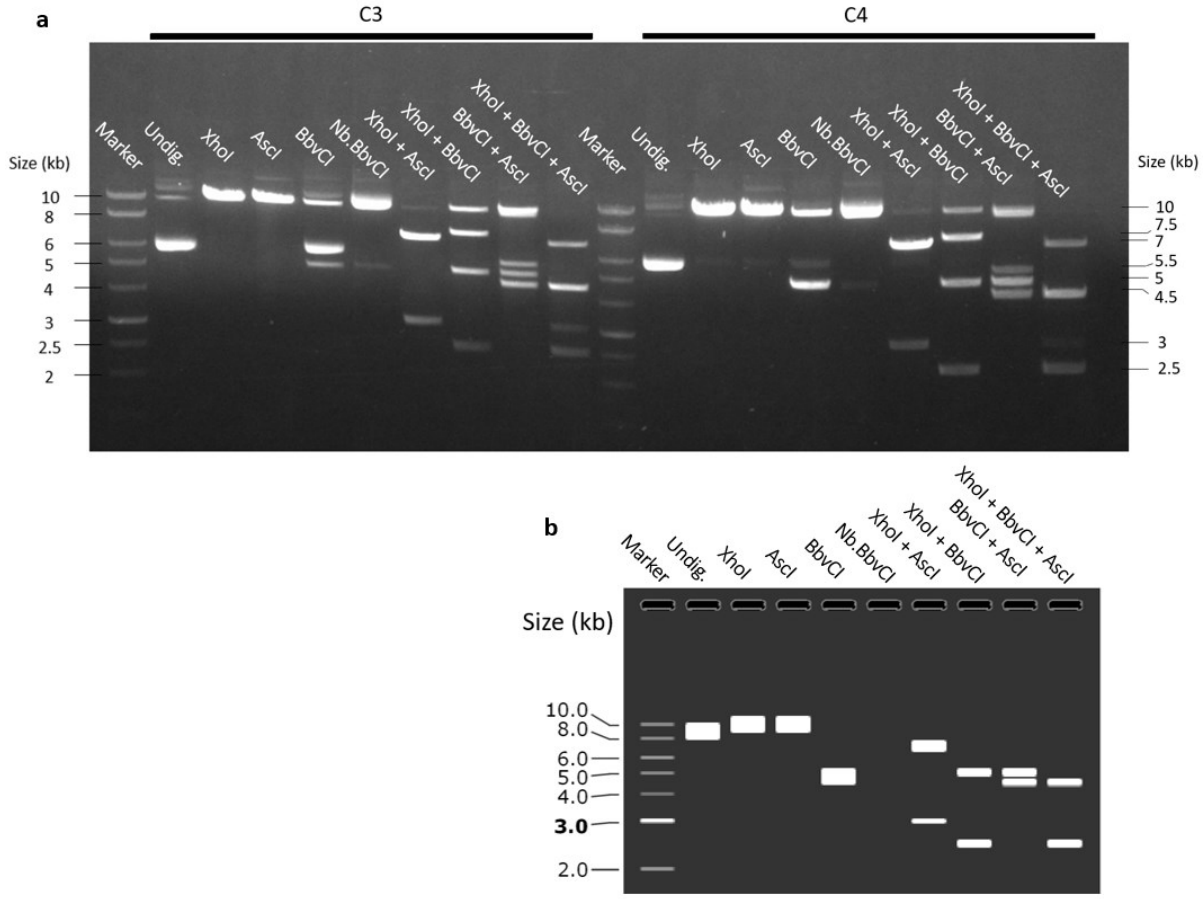


Figure 11. *(**a**) Diagnostic gel obtained after digestion of clones C3 and C4 by different combinations of the XhoI, BbvCI, and AscI restriction enzymes. The result of a digestion with Nb.BbvCI and an undigested negative control are also present. (**b**) Gel simulation of the expected band patterns, assuming the digestion reactions run till completion. It was not possible to simulate the nicking reaction with Nb.BbvCI and therefore this lane is left empty. The simulation was made using SnapGene®5.2.4.*

R.1.2.3 Clone 4 likely contains the correct plasmid

The results obtained from the Nb.BbvCI digestion of clone 4 (**Fig. 11a, lane 16**) show an additional band corresponding to fragments of 5 kb. These fragments are likely caused by star activity of the Nb.BbvCI enzyme, as previously described. **Lane 18** (**Fig. 11a**) contains an unexpected band of about 7.5 kb and one band corresponding to the open linear conformation. The band of 7.5 kb can be explained by partial cleavage of the plasmid at the XhoI site and a BbvCI site in one of the two BbvCI clusters. This would result in fragments of 2.5 kb and 7.5 kb, which could explain the presence of the

7.5 kb band. In **lane 19** (**Fig. 11a**), there seems to be an additional band of 5.5 kb and one corresponding to the open linear conformation of the plasmid. The 5.5 kb band can be explained by partial cleavage of the plasmid on the AscI site and a site in the BbvCI cluster most distant from the AscI cut site (**Fig. 9**). This combination would result in a fragment of 4.5 kb and a fragment of 5.5 kb, explaining the presence of the 5.5 kb band. **Lane 20** (**Fig. 11a**) shows two unexpected bands corresponding to the open linear conformation of the plasmid and a band of 7 kb. The band of 7 kb could be explained by partial cleavage of the plasmid on the XhoI site and the AscI site. This would result in fragments of 7 kb and 3 kb, both of which are present. To conclude, all bands are as expected or can be accounted for. Therefore, we conclude that clone 4 also contains the desired plasmid.

R.1.3 DNA sequencing

To further characterize the plasmids obtained from clones 3 and 4, the overlapping regions from the gBlocks were sequenced. The rationale behind this is as following: if all sequences of the overlapping regions agree with the expected sequence, then the gBlocks should be assembled in the correct order. The gBlocks are sequenced verified by the company from which they were ordered, therefore we conclude that the plasmids have the desired sequence if the gBlocks are assembled in the correct order. Additionally, the key regions were sequenced, which include: The two HtH sites, the Nb.BbvCI nicking sites (i.e. both fork insertion regions), the XhoI cut site (i.e. the biotinylation site) and the AscI cut site. Because these regions are already sequenced extensively these results serve only as an additional confirmation of what is already known.

R.1.3.1 Plasmids purified from clones 3 and 4 can be used to synthesize the construct

The sequencing results from clone 3 (**Fig. 12a**) agreed well with expectations. All the overlapping regions matched the expectation. Furthermore, all the key regions were sequenced and matched the expectation as well, except for one of the HtH sites. Sequencing samples for this HtH site were sent multiple times. However, none of them agreed with the expected sequence. All sequencing results from this region seem to be degraded and covered an almost identical short region, indicating that most likely the working stock of the sequencing primer for this region was contaminated, resulting in degraded and erroneous sequencing results.

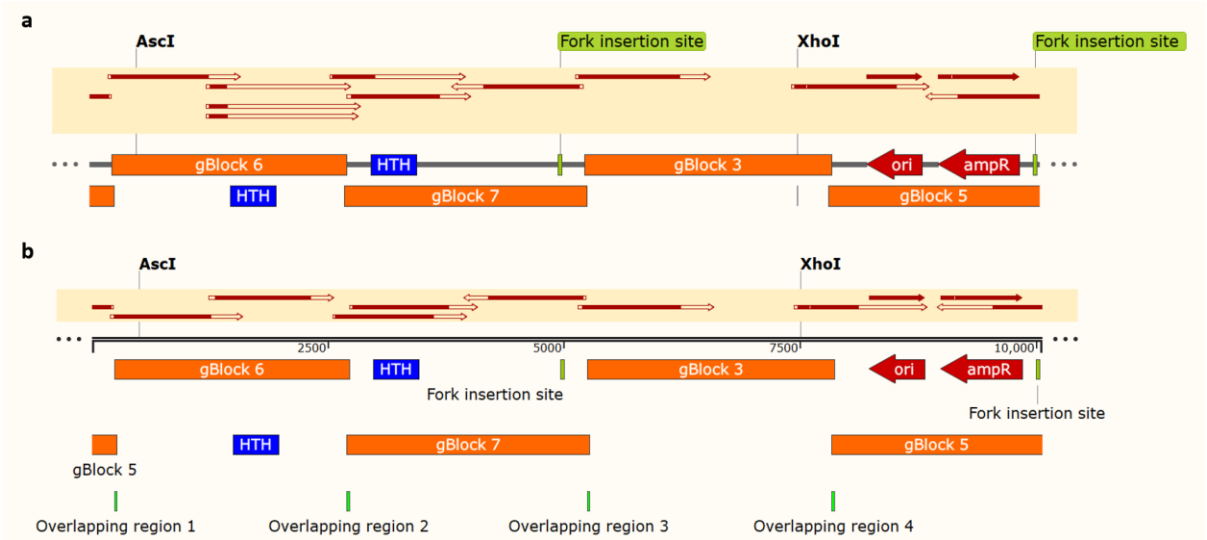


Figure 12. *Plasmid maps showing the different plasmid features and sequencing results. Results are represented by red/yellow arrows, the red regions correspond to successfully sequenced and matching sequencing reads. (**a**) Sequencing results and plasmid map from clone 3. (**b**) Sequencing results and plasmid map from clone 4. The figure was made using SnapGene®5.2.4.*

Furthermore, the results of the sequencing reactions from clone 4 (**Fig. 12b**) agree with the expectations as well. All overlapping regions were present and corresponded to the expectation. Moreover, all key regions were sequenced successfully and matched the expectation.

Combining all biochemical evidence together with the sequencing results, we conclude that the plasmids purified from both clones have the desired sequence, the correct order of gBlocks, and contain all the desired features and can therefore by used to synthesize the construct.

R.2 Construct synthesis and characterization

R.2.1 Construct synthesis and biochemical characterization

The purified plasmids from clone 4 were chosen to synthesize (**see M.1**) the construct for the single-molecule experiments. For this first batch of constructs, the nicking reaction was carried out for 3 h and 5' phosphorylated oligonucleotides containing a Cy5 fluorophore at their 3' end were incorporated during fork annealing. At each step, except after nicking, a sample was taken from the batch which was subsequently used to run a diagnostic gel to monitor the synthesis process. Furthermore, a digestion reaction of the final construct by AscI was conducted for 55 min at 37 °C to see whether both forks were incorporated correctly. The samples were run on a 0.8% agarose gel in TAE using ethidium bromide to stain the DNA.

R.2.1.1 Expected band patterns on diagnostic gel

The band patterns obtained from the samples taken after each step of the synthesis process are expected to yield a single band of 10 kb, corresponding to the open linear conformation. Once the plasmids are linearized by XhoI, the subsequent steps do not alter the state of the DNA significantly enough to be visible on the gel. Furthermore, the samples taken after fork annealing, ligation, and digestion by AscI are expected to be visible when imaging Cy5 fluorescence, whereas the undigested, linearized, and biotinylated plasmids are not expected to emit a fluorescence signal. Digestion of the construct with AscI is expected to yield two fragments of 3 and 7 kb. Since both fragments should be present in equimolar amounts and the 7 kb fragments contain a factor 2.33 as many base pairs as the 3 kb fragments, the 7 kb band is expected to be a factor 2.33 as intense when imaging ethidium bromide. In contrast, the intensity profile of these bands should be equivalent when imaging fluorescence (Cy5), because the AscI cut site is located such that both fragments should contain one fluorescent fork, as shown in the schematic representation of the desired construct (**Fig. 13c**).

R.2.1.2 The diagnostic gel suggests the desired construct was synthesized successfully using the plasmids obtained from clone 4

The most intense bands from the experimental results (**Fig. 13a-i**) correspond to the expected band patterns. However, in the last three lanes additional unexpected bands are present as well. In **lane 5 and 6** (**Fig. 13a-i**), there are additional bands corresponding to fragments of 2.5 kb, 5 kb, and 7.5 kb, which can be accounted for by star activity of the Nb.BbvCI enzymes. Digestion of the final construct

by AscI also produced additional unexpected band patterns (**Fig. 13a-i, lane 7**), corresponding to fragments of 10 kb linearized, 0.5 kb, 2.5 kb, and 4.5 kb. The 10 kb band corresponding to the open linear conformation of the plasmid is due to incomplete digestion, the 4.5 kb and the 0.5 kb fragments can be explained by cleavage of the undesired fragments which were produced by Nb.BbvCI during the previous step of the synthesis process. Cleavage of the 5 kb fragments by AscI yields fragments of 4.5 and 0.5 kb. Furthermore, the 2.5 kb fragments seem to be the same fragments of 2.5 kb of the previous step further supporting this. As expected, the lanes corresponding to the undigested sample and the samples taken after linearization and biotinylation do not emit any fluorescence signal, whereas those obtained after fork annealing, ligation, and AscI digestion do (**Fig. 13b-i**). Moreover, after purification of the construct (**see M.1.6**), the final yield was calculated to be 13.9%. Taken together, the experimental data shows that the construct seems to be synthesized successfully from the plasmids obtained from clone 4.

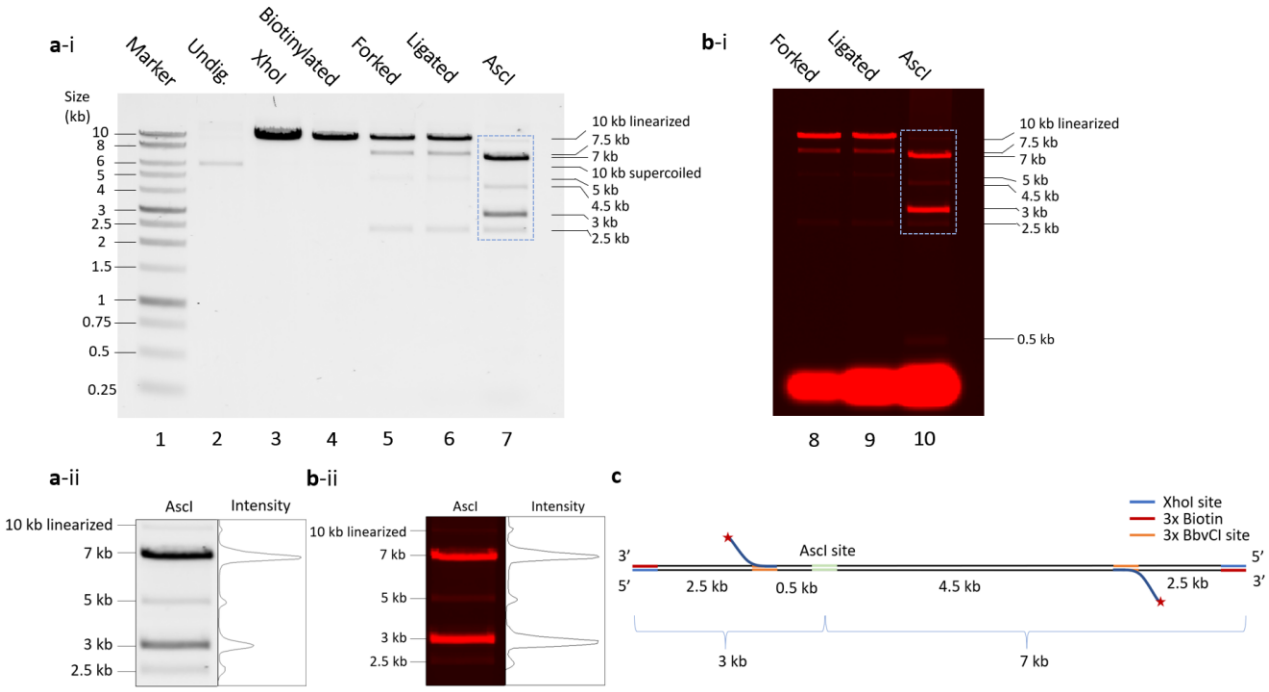


***Figure 13**. Visualizations of the diagnostic gel, all gel images were taken from the same gel. (**a-i**) Visualization of ethidium bromide-stained DNA. (**a-ii**) Intensity profile of the dashed region from **a-i**, the ratio of the areas under the curve of the bands corresponding to 7 kb and 3 kb was calculated to be 2.26. (**b-i**) Visualization of fluorescence signal (Cy5) from the last three lanes of **a-i**, the large red spots at the bottom of the gel are unincorporated fork oligonucleotides which were added in excess during the annealing process. (**b-ii**) Intensity profile of the dashed region from **b-i**, the ratio of the areas under the curve of the bands corresponding to 7 kb and 3 kb was calculated to be 0.95. Lane 1 corresponds to the DNA ladder. Lane 2 shows the undigested plasmid, lanes 3-6 correspond to samples taken after each step in the synthesis process of the construct and in lane 7 the fragments obtained after digestion with AscI are shown. Lanes 8-10 are the same as lanes 5-6 but showing the Cy5 signal instead (**c**) Schematic representation of the construct showing the position of the AscI site in lime green. Intensity profiles were obtained using FIJI2.1.0.*

R.2.1.3 Fork oligonucleotides are incorporated with approximately equal efficiencies at either fork insertion site.

In order to roughly assess the fork incorporation efficiency, densitometry measurements were taken from the gel images. The ratio of the area under the curve of the bands corresponding to 7 kb and 3 kb obtained after digestion by AscI and visualisation of ethidium bromide-stained DNA was calculated to be 2.26 (**Fig. 13a-ii**), agreeing well with the expected value of 2.33. The same ratio was found to be 0.95 from the fluorescence signal (**Fig. 13b-ii**), which agrees well with the expected value of 1. Therefore, from the digestion with AscI, it appears that the forks were incorporated with approximately equal efficiency at either fork insertion site.

R.2.1.4 Unexpected band patterns in diagnostic gel are caused by star activity of Nb.BbvCI

To assess whether the unexpected fragments present in the last three lanes of the diagnostic gel (**Fig. 13a-i**) were indeed due to star activity, the digestion was repeated after the construct was purified (**see M.1.6**). In this reaction, 350 ng of purified construct was digested by 5 units of AscI for 2 h at 37 °C and run on a 0.8% agarose gel in TAE using ethidium bromide to stain the DNA (**Fig. 14a**). Besides a band corresponding to the open linear conformation, two bands of 3 kb and 7 kb are present, as expected. This time, however, the 2.5 kb, 5 kb, and 7.5 kb bands are not present, meaning these fragments did not result from cleavage by AscI and are likely due to star activity of the Nb.BbvCI enzyme. Moreover, as expected both the undigested and digested construct samples show fluorescence (**Fig. 14b**). There are also faint smears present which are due to degradation products caused by contamination of the samples by DNases. The DNases also caused the undigested plasmid to be linearized. To conclude, the previously observed fragments of 2.5 kb, 5 kb, and 7.5 kb are likely due to star activity of the Nb.BbvCI enzyme, therefore the construct seems to be synthesized successfully and all the obtained bands from the diagnostic gel (**Fig. 13a-i**) are as expected or can be accounted for.

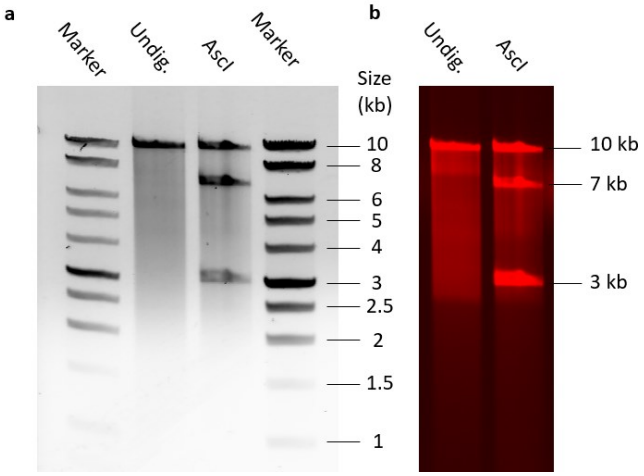


***Figure 14**. Visualizations of the purified construct digested by AscI, both images were taken from the same gel. (**a**) Visualization of ethidium bromide-stained DNA. (**b**) Visualization of fluorescence signal (Cy5), shown only for the lanes corresponding to the undigested sample and the purified construct digested by AscI.*

R.2.2 Biophysical characterization of the construct

After synthesis and biochemical characterization of the construct, the final test was to biophysically characterize it by correlative dual beam optical trapping in combination with confocal fluorescence microscopy. To this end, approximately 3.6 pM of construct DNA was dispensed into the second channel of the C-trap during a single-molecule experiment. In total, 49 constructs were optically

trapped and analyzed. All fluorescence scans were obtained at a 2 pN force-clamp, to image the trapped molecules under similar extension. Furthermore, during imaging, the beads were not moved in the X- or Y-dimension. The fluorescence scans obtained in this manner, were used to assess whether two fluorescent forks are inserted correctly and to determine the fork incorporation efficiency. Furthermore, the force-extension data obtained from this experiment was fitted to an extensible and inextensible worm-like chain model (**see T.2**), to this end Matlab scripts which were already present in the laboratory were converted and adapted to Python 3.8 to help the laboratory streamline the data analysis. Lastly, the end-to-end DNA length was determined by brightfield microscopy.

R.2.2.1 Fitted values from inextensible and extensible worm-like chain models agree well with each other but not with the expectations.

To obtain the WLC model fit, a subset of the data where the force ranges from 1-10 pN and the extension from 0-3.5 µm was used (**Fig. 15a-i**). The extension was filtered in addition to the force, because sometimes the optically trapped DNA molecules broke at large extensions, causing forces in the range 2-10 pN to be detected at extensions above 3.5 µm. When DNA molecules break the force needed to extend the DNA suddenly drops to zero, during this drop the software sometimes detects forces between zero and the force exerted just before the drop.

The canonical value for the persistence length is approximately 50 nm, however, the persistence length is dependent on the ionic strength of the medium in which the DNA molecules are imaged. The ionic strength of the imaging buffer, which contains 50 mM potassium glutamate and 10 mM magnesium acetate, was calculated to be 80 mM using **equation 3** taken from Brunet et al., 2015, where one half is because we are including both cations and anions, c_i is the molar concentration of ion i, z_i is the charge number of that ion, and the sum is taken over all ions in the solution. Based on what others have found and the calculated value for the ionic strength of the medium in which the force-extension data was obtained, the persistence length is expected to fall within the range of 33-50 nm[32,31,37]. Furthermore, the expected value for the contour length is 3448 nm, which is calculated by multiplying the number of base pairs, 10,052 bp, by the length per base pair of B-form DNA, which is approximately 0.343 nm. **Figure 15a-i** shows two purple dashed lines which indicate the boundaries of a region in which the force-extension data is expected to fall, based on the expected values for the persistence length and the contour length.

$$\boldsymbol{equation\ 3.} \qquad \frac{1}{2}\sum_{i=1}^{n} c_i z_i^2$$

The values obtained from a two-parameter fit of the persistence length and the contour length to the force-extension data are (Mean ± SEM) 20 ± 5 nm and 3432 ± 81 nm, respectively (**Table 3**). The value for the contour length obtained from the fit agrees well with the expected contour length of 3448 nm. Furthermore, the fit is in good agreement with the experimental data, which can be seen by looking at **Figure 15a-ii** where the residuals are plotted. The magnitude of the mean residual value was calculated to be 16 nm over 2855 datapoints, which is close to zero, meaning that most of the residuals are likely caused by random noise. However, the value for the persistence length is rather low, causing the measured data to fall outside of the expected region. The deviation from the expected outcome does not seem to originate from the fitting algorithm, because the fit seems to be in good agreement with the experimental data. Our best guess is that the deviation from the expectation results from the diffraction-limited (~350 nm) resolution of the brightfield microscope, which is used to measure the extension of the trapped molecules. It would be interesting to redo these measurements during a

magnetic tweezers assay, which is known to be able to do more sensitive measurements at low forces, to see whether it results in similar force-extension data.

Model	Persistence length (nm)	Contour length (nm)	Elastic stretch modulus (pN)	|Mean residual value|	Force range (pN)
WLC	20 ± 5	3432 ± 81	-	16	1-10
xWLC	18 ± 3	3449 ± 30	2730 ± 1125	1	1-50

Table 3 *. Values obtained from fitting the inextensible and extensible worm-like chain models to the force extension data in the specified force range (Mean* ± SEM*).*

To fit the extensible worm-like chain model a subset of the data where the force ranges from 1–50 pN and the extension from 0-3.5 µm was used (**Fig. 15b-i**). The canonical value for the elastic stretch modulus of DNA is approximately 1000 pN, however its value is also influenced by the ionic strength of the medium. In contrast to the persistence length, the elastic stretch modulus increases with increasing ionic strength[32]. Based on what others have found and the calculated value for the ionic strength of the medium in which the force-extension data was obtained, the elastic stretch modulus is expected to fall within the range of 1000-1435 nm[32,31,37]. **Figure 15b-i** shows two purple dashed lines which indicate the boundaries of a region in which the force-extension data is expected to fall, based on the expected values for the persistence length, elastic stretch modulus, and the contour length.

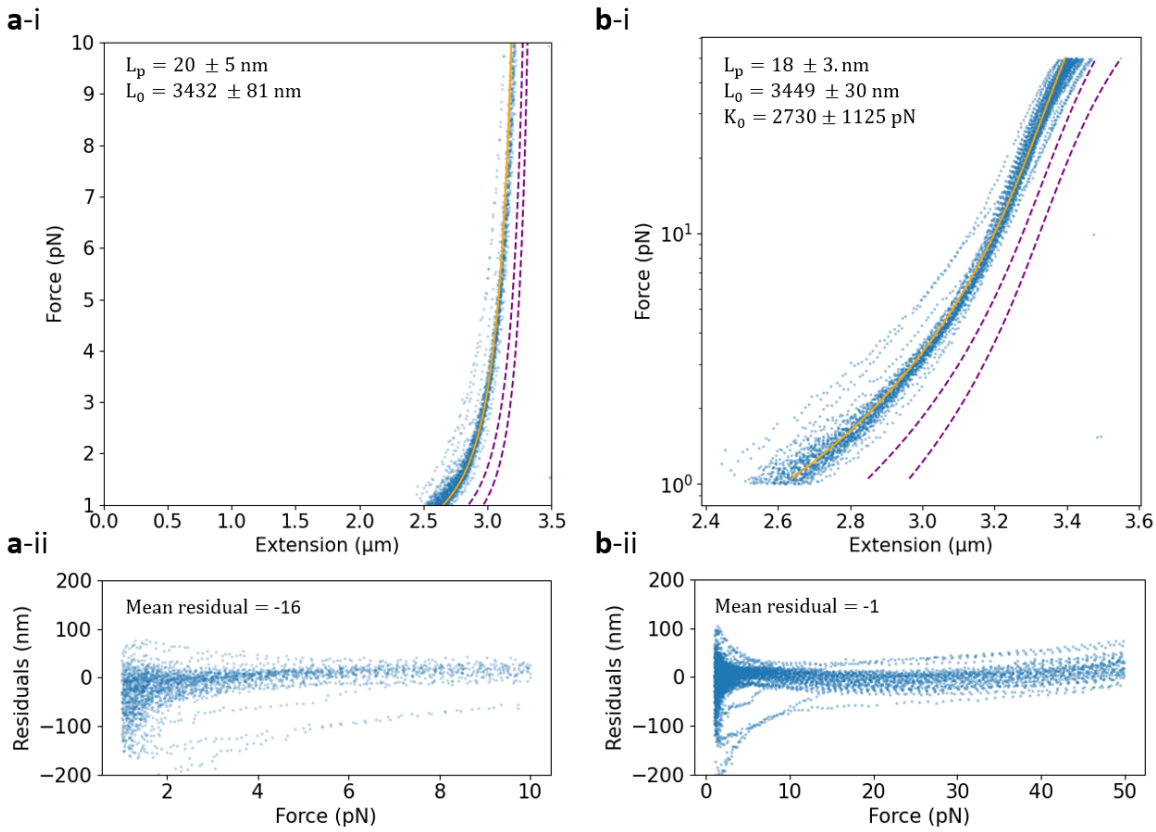


Figure 15*. Force-extension data obtained from single-molecule experiments, N=49.****(a-i)****; (****Blue****) Experimental data. (****Orange****) Inextensible worm like chain fit, to obtain the fit a subset of the data was used where the force ranges from 1–10 pN and the extension from 0-3.5 µm. (****Dashed purple****) The data*

is expected to fall in the region bounded by these lines. ***(b-i)****; (****Blue****) Experimental data. (****Orange****) Extensible worm like chain fit, to obtain the fit a subset of the data was used where the force ranges from 1–50 pN and the extension from 0-3.5 µm. The forces are plotted on a logarithmic scale with base 10 for clarity. (****Dashed purple****) The data is expected to fall in the region bounded by these lines.* ***(a-ii)****; Shows the residuals versus force behavior for the WLC model fit. (****b-ii****) Shows the residuals versus force behavior for the xWLC model fit. Fitted values are mean ± std.*

A three-parameter fit resulted in values for the persistence length, contour length and the elastic stretch modulus of (mean ± std) 18 ± 3 nm, 3449 ± 30 nm, and 2730 ± 1125 pN, respectively (**Table 3**). The obtained value for the contour length is again in good agreement with the expected value. Furthermore, the calculated value for the persistence length agrees well with the fitted value from the WLC model. The stretch modulus is higher than expected, however there is a relatively large uncertainty. The magnitude of the mean residual value (**Fig. 15b-ii**) is calculated to be 1 over 4995 datapoints, which is again close to zero, showing the model provides a decent fit to the data. As before, the experimental data falls outside of the expected region which is likely because of the same reasons as stated above.

For both models, the deviation from the mean residual value increases at the two extremes of the force range. This behavior at low force is likely due to increasing inaccuracies of the measurement apparatus. Whereas at high forces, the DNA starts to yield cooperatively, causing the force-extension curve to plateau, explaining why the mean residual value starts to deviate more and more at high forces. Additionally, by brightfield microscopy the mean end-to-end distance of the optically trapped constructs was calculated to be 2.85 ± 0.03 µm when the molecules are held at a 2 pN force-clamp. This was calculated by calculating the mean distance between the center of the beads and subtracting the bead diameter. Furthermore, this value is in good agreement with the mean extension calculated from the force-extension data at 2 pN which is calculated to be (Mean ± SEM) 2.87 ± 0.02 µm.

R.2.2.2 Two forks can be incorporated successfully

The fluorescent scans obtained from the single-molecule experiments were used to determine the fork incorporation efficiency. The results are shown in **Table 4**. As expected some of the optically trapped construct molecules showed two red foci, corresponding to the fluorescent fork oligonucleotides. **Figure 16a** shows a schematic representation of the construct and **Figure 16b** shows the average intensity profile from the red (Cy5) color channel of an exemplary scan, where two foci corresponding to the fork oligonucleotides are present. The percentage of constructs that incorporated both forks was calculated to be 35.8%. This value is likely a slight underestimation because some fluorophores may have photobleached before imaging. Taken together, this data shows the fluorescent forks are successfully incorporated during the synthesis process and about one in three trapped constructs incorporated both forks.

Number of incorporated forks	Number of observations	Occurrence (%)
0	12	24.5
1	19	38.8
2	18	36.7

Table 4*. Fork incorporation statistics as determined by fluorescence microscopy, N=49.*

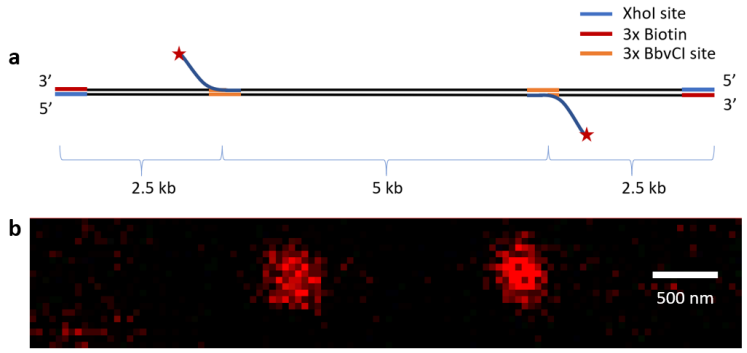


***Figure 16**. Exemplary data, showing a construct which incorporated both forks. (**a**) Cartoon of the construct. (**b**) Average intensity over all frames of the red color channel showing two foci corresponding to the Cy5 labeled forks. Scale bar applies to both scans.*

R.3 Construct synthesis according to adjusted protocol and further biophysical characterization

R.3.1 Construct synthesis

To further biophysically characterize the construct, a new batch was produced using a slightly altered synthesis process. Instead of using Cy5 labeled fork oligonucleotides, Cy3 was used to fluorescently label the fork oligonucleotides at their 3' ends during the synthesis process (**see M.1.4**). The Cy3 fluorophore is expected to take longer to photo bleach, increasing the accuracy with which the position of the forks is determined. Furthermore, instead of nicking for 3 h the nicking reaction was conducted for 5 h (**see M.1.3**), in the hope of enhancing the fork incorporation efficiency. Increasing the time during which the DNA is nicked should in theory result in a higher fork incorporation efficiency. However, it is also expected to decrease the final yield, because the increased nicking time results in increased star activity as well. Therefore, a trade-off needs to be made between higher fork incorporation efficiency versus an increased final yield. The final yield previously obtained was so high it had to be diluted before it could be used during the single-molecule experiments. Additionally, the DNA constructs cannot be frozen and can at most be used for 1 month. In our case, a single synthesis generates enough DNA for several weeks. Thus, a higher fork incorporation efficiency is preferred over a higher final yield.

R.3.1.1 The construct can also be synthesized successfully from plasmids obtained from Clone 4 when nicking for 5 h.

Except for the alterations mentioned above, the exact same protocol was followed as described in section **M.1** during the synthesis process. To monitor the synthesis process samples were taken after each step. This time we also included a sample which was taken after the nicking reaction to make for a more complete diagnostic gel and to support the claim that the additional bands previously observed (**see R.2.1.4**) are due to star activity. The expected band patterns are the same as described in section **R.2.1.1**. Furthermore, the additional sample taken after the nicking reaction is expected to result in fragments of 10 kb on the gel, besides the bands which are expected to result from star activity.

The experimental results (**Fig. 17**) agree well with the expectation. As expected, the additional bands of 2.5, 5, and 7.5 kb arose after the nicking reaction (**Fig. 17, Lane 5**), supporting the claim that these are due to star activity of the Nb.BbvCI enzyme. However, some lanes include a faint smear which is most likely due to DNases, which were unintendedly introduced and degraded some of the DNA. Furthermore, the longer nicking time did not result in a significant decrease of the final yield, which was calculated to be 13.9% for the previous batch of constructs and 12.8% for this batch of constructs

after purification (**see M.1.6**). To conclude, nicking for 5 h instead of 3 h also results in the successful synthesis of the desired construct.

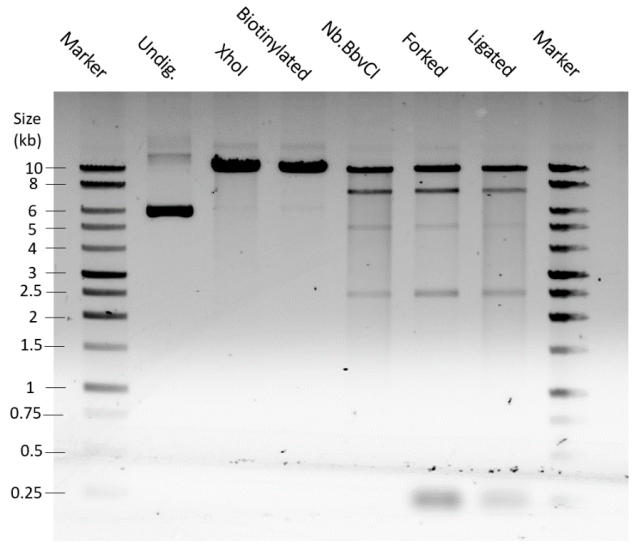


***Figure 17**. Visualization of the ethidium bromide-stained DNA from the diagnostic gel. The first and last lane correspond to the DNA ladder. Lane 2 corresponds to the undigested plasmid DNA. Lanes 3-7 correspond to samples taken after each step of the construct synthesis process.*

In contrast to the diagnostic gel used to monitor the synthesis process of the previous batch of constructs (**see R.2.1**), this diagnostic gel did not include a sample taken after digestion of the final construct by AscI, because it is not possible to measure the fluorescence signal obtained from Cy3 separately from the signal obtained by ethidium bromide. In this case, the only way to determine the fork incorporation efficiency is by confocal fluorescence microscopy during single-molecule experiments.

R.3.2 Further biophysical characterization of the construct

After successful synthesis of the second batch of constructs, we wanted to determine the fork incorporation efficiency and compare these values to those obtained from the previous batch of constructs, to assert whether the altered protocol resulted in an enhanced fork incorporation efficiency as expected. Furthermore, the mean lifetime of the Cy5 and Cy3 labeled fork oligonucleotides were determined, to confirm the Cy3 fluorophore takes longer to photobleach than the Cy5 fluorophore.

Moreover, fluorescently labeled dCas9-Halo-AF488 and dCas9-Halo-JF646 complexes were loaded in bulk on the constructs (**see M.2.1**) to assess whether these can be used to determine the orientation of single optically trapped construct molecules. To answer these questions, single construct molecules from the second batch were optically trapped and analyzed by correlative dual beam optical trapping in combination with confocal fluorescence microscopy. In total, 43 constructs were optically trapped and analyzed.

In addition, as a final test, we wanted to assert whether MCMs could be loaded successfully onto the construct and form DHs at the artificial origin of replication sites using the reconstituted biochemical pathway. To this end, [JF646-Halo-Mcm3]MCM rings were loaded in bulk by ORC, Cdc6, and Cdt1 (**See M.2.2**). To assert succesful loading, another single-molecule experiment was conducted during which single construct molecules, again from the second batch, were optically trapped and analyzed by correlative dual beam optical trapping in combination with confocal fluorescence microscopy. From this experiment, a total of 13 constructs were optically trapped and analyzed.

During both single-molecule experiments approximately 5 pM of construct DNA was dispensed into the second channel of the C-trap and the fluorescence scans were taken at a 2 pN force-clamp. Furthermore, during imaging, the beads were not moved in the X- or Y-dimension. The fluorescent fork oligonucleotides are used as internal nanorulers to determine the distances to the center of the constructs for all in bulk loaded proteins and compare these distances to the expectations to confirm succesful loading. In this section the results from these two single-molecule experiments will be presented and discussed in order.

R.3.2.1 Adjusted synthesis protocol results in a significantly enhanced fork incorporation efficiency.

The fork incorporation efficiency of the second batch is more than twice as large than that of the previous batch. Of all trapped constructs 83.7% (**Table 5**) incorporated both forks, whereas the fork incorporation efficiency obtained from the previous batch of constructs was only 36.7% (**Table 4**). Moreover, not a single construct molecule was trapped which incorporate zero fluorescent forks during this experiment. The main difference between the synthesis processes of these two batches is the nicking time, which was increased from 3 to 5 h during synthesis of the second batch. To conclude, nicking for 5 instead of 3 h during construct synthesis significantly enhances the fork incorporation efficiency, as expected.

Number of incorporated forks	Number of observations	Occurrence (%)
0	0	0
1	7	16.3
2	36	83.7

***Table 5**. Fork incorporation statistics as determined by fluorescence microscopy, N=43.*

R.3.2.2 Cy3 fluorophore takes considerably longer to photobleach compared to the Cy5 fluorophore

In addition, the fluorescence scans obtained from this dataset were used to determine the mean lifetime of the Cy3 fluorophore and compare it to the mean lifetime of the Cy5 fluorophore. The mean lifetime of the Cy5 fluorophore was determined using the data from the single-molecule experiments from the first batch of constructs. The mean lifetime of Cy5 and Cy3 were calculated as stated in section **M.5.3** and are 3.62 ± 0.07 (**Fig. 18a**) and 68.3 ± 0.7 (**Fig. 18b**) frames, respectively. The duration of a frame is equivalent to 0.6 s.

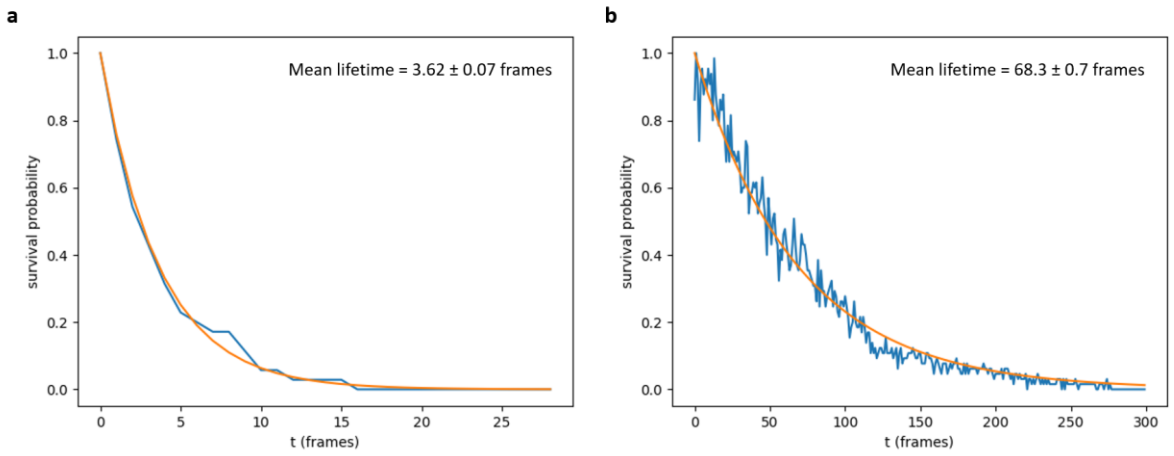


***Figure 18**. Survival probabilities over time of the Cy5 (red) and Cy3 (green) fluorophores. (**a**); (**Blue**) Experimental data obtained from Cy5 fluorophores, N=49. (**Orange**) Fitted exponentially decaying function (**b**); (**Blue**) Experimental data obtained from Cy3 fluorophores, N=43. (**Orange**) Fitted exponentially decaying function.*

The survival probability curves (**Fig. 18**) are expected to monotonically decrease over time, however the curves sometimes rise and then fall again, which we account to the detection of false positives and errors in foci detection. Therefore, the mean lifetimes are a slight overestimation of the actual values. Taken together, the data show the Cy5 fluorophore bleaches considerably faster than the Cy3 fluorophore, as expected.

R.3.2.3 The dCas9 complexes can be used to determine the orientation of single optically trapped construct molecules

Lastly, we used this data set to assert whether the in bulk loaded dCas9 complexes can be used to determine the orientation of single optically trapped construct molecules, which would be useful for future experiments. The dCas9-Halo-AF488 and dCas9-Halo-JF646 complexes were incubated with gRNAs DRM 1 and 2 (**See appendix**), respectively, and are expected to bind on opposing off-center regions at the ends on the construct molecules.

Furthermore, to assess whether these complexes bound at the expected distance from the construct center, a Gaussian mixture model was fitted to the raw localization distribution of the ensemble of detected foci, corresponding to the dCas9 complexes, over all fluorescence scans from all optically traped construct molecules. These distances were then expressed relative to the mean center of the constructs as described in section **M.5.2**. Per color channel, the distances to the center are then averaged to get a mean value for the distance to the center.

The dCas9-Halo-AF488 and dCas9-Halo-JF646 are expected to bind 3.2 and 3.0 kb from the center of the construct, respectively, where the middle of their gRNA binding sites was taken as the expectation. Using the biochemical knowledge that the distance between the fork insertion sites is 5 kb, a micron-to-kilobase-pair coordinate conversion factor can be calculated (**see M.5.2**). In this manner, the observed distances of the dCas9 complexes relative to the mean center can be expressed in kb and compared to the expectations.

As expected, on all optically trapped contruct molecules, the dCas9-Halo-JF646 complex bound at the opposing end relative to the dCas9-Halo-AF488 complex. **Figure 19** shows averaged intensity scans from two exemplary optically trapped constructs, where **Figure 19a** shows a schematic of the two different orientations in which the construct can be trapped, **Figure 19b** shows the signal obtained from the fork oligonucleotides, and **Figure 19c and d** show the signals obtained from the dCas9 complexes, which bind to opposing ends on the construct.

Furthermore, using the fluorecent fork oligonucleotides as internal nanorulers the distances of the foci corresponding to dCas9-Halo-JF646 and dCas9-Halo-AF488 to the mean center of the constructs were determined and compared to the expectation. From the calculated conversion factor the base pair length was calculated to be (Mean ± SEM) $0.357 \pm 1.7 * 10^{-4}$ nm. **Figure 20a** shows the positional distributions and a fitted Guassian mixture model of the color channel corresponding to the fluorescent fork oligonucleotides (Cy3) and **Figure 20b and c** show the localisation distributions of

dCas9-Halo-JF646 and dCas9-Halo-AF488, respectively. All positions are expressed relative to the mean center of the constructs.

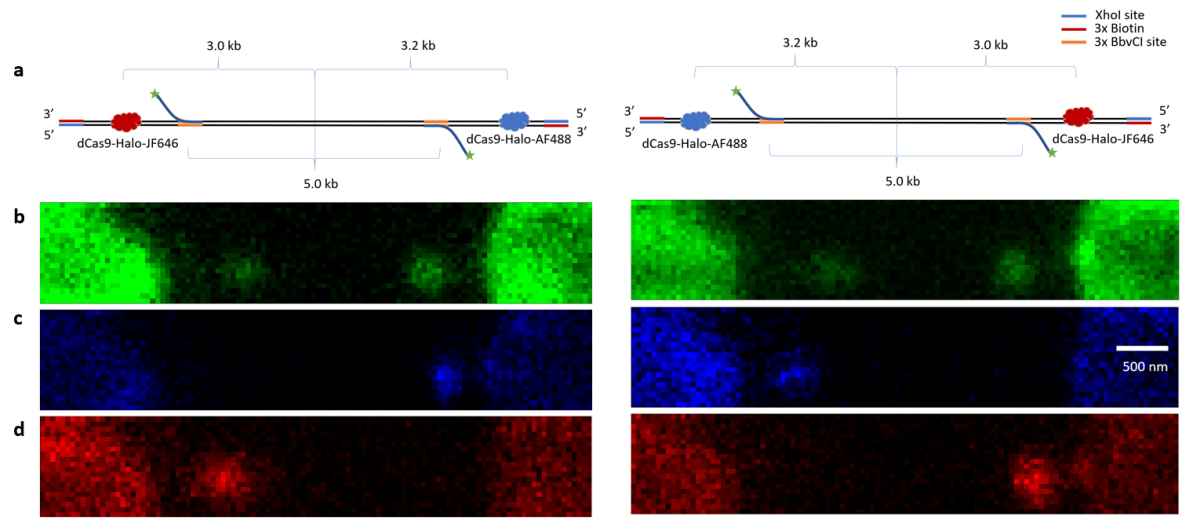


***Figure 19**. Averaged intensity per color channel of two exemplary construct molecules which were trapped in different orientations. (**a**) Schematic representation of construct orientations. (**b**) Green color channel showing the Cy3 labeled forks. (**c**) Blue color channel, showing dCas9-Halo-AF488. (**d**) Red color channel, showing dCas9-Halo-JF646. Scale bar applies to all fluorescence scans. The field of view also includes half of both optically trapped polysterene beads at the edges. Averaged intensity profiles were obtained using FIJI2.1.0.*

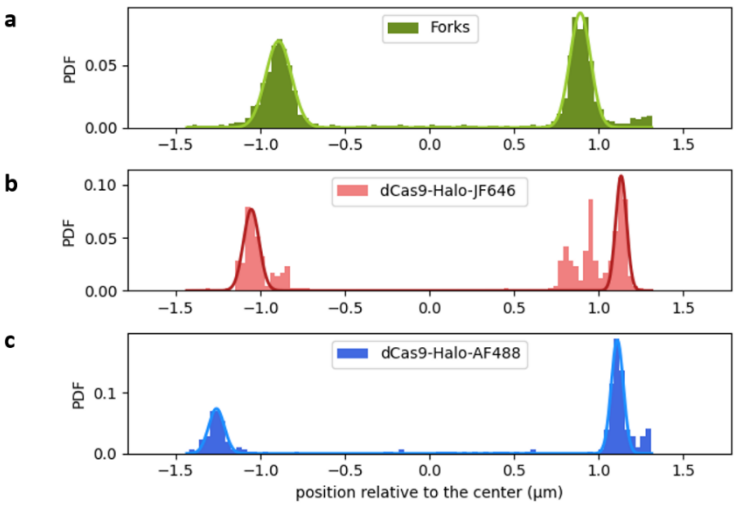


***Figure 20**. Localization distributions from the ensemble of all detected foci over all trapped constructs per color channel realtive to the mean center of the constructs. (**a**) The center between the fitted means of the peaks corresponding to the forks was identified as the center of the construct. (**b**) Distances to the center of red fluorescent dCas9-Halo-JF646. (**c**) Distances to the center of blue fluorescent dCas9-Halo-AF488.*

The mean distances to the center (**Table 6**) for the dCas9-Halo-JF646 and dCas9-Halo-AF488 were found to differ from the expectation by 13 and 38 nm, respectively, showing that the dCas9 complexes not only bind at opposing ends, but also at the expected positions on the construct relative to the center. However, some of the dCas9-Halo-JF646 complexes were observed to bind off-target to other

regions on the construct. Therefore, it might be interesting to optimize the gRNA sequences to decrease off-target binding and enhance the binding affinity to the desired location.

Protein	Distance from center (μm)	Distance from center (kb)	Expected distance from center (kb)
dCas9-Halo-JF646	$1.1 \pm 8.0 * 10^{-3}$	$3.1 \pm 2.2 * 10^{-2}$	3.0
dCas9-Halo-AF488	$1.2 \pm 4.8 * 10^{-3}$	$3.3 \pm 1.3 * 10^{-2}$	3.2

***Table 6**. Observed (Mean ± SEM) and expected distances from the center of the dCas9 complexes, N=43.*

Taken together, this data shows that using fluorescent DNA oligos as internal nanorulers allows us to determine the positions of the fluorescent dCas9 complexes with sub-pixel precision (<50 nm) and shows both gRNA sequences, together with the fluorescent dCas9 complexes, can be used to determine the orientation of single optically trapped construct molecules.

R.3.2.4 MCMs and DHs are successfully loaded onto the construct at the artificial origin of replication sites

As a final test, we wanted to assert whether MCMs could be loaded successfully onto the construct and form DHs at the artificial origin of replication sites using the reconstituted biochemical pathway. To this end, [JF646-Halo-Mcm3]MCM rings were loaded in bulk by ORC, Cdc6, and Cdt1 (**See M.2.2**) and analyzed during a second single-molecule experiment using construct molecules from the second batch.

To assess whether these complexes bound at the expected distance from the center, a Gaussian mixture model was again fitted to the raw localization distribution of the ensemble of detected foci corresponding to the MCM complexes over all scans from all optically traped construct molecules. We defined the expected distance to the center of the MCMs as the mean of the differences to the center of the middle between the forward and reverse ORC binding sites, which is approximately 0.9 kb. We defined it as such, because both inter ORC binding sites are not exactly the same distance from the center. To ascertain the orientation of the DNA, we also bound dCas9-Halo-AF488/gRNA DRM 1 RNP complexes in bulk, which we previously showed worked well (**see R.3.2.3**). The expected binding site of dCas9-Halo-AF488 is the same as stated previously.

From the experimental data (N=13) we observed red foci, corresponding to [JF646-Halo-Mcm3]MCMs, showed two bleaching steps of approximately equal size, showing DHs can be successfully loaded onto the synthesized DNA. **Figure 21a** shows an exemplary trapped construct, a schematic representation of the orientation is shown in **Figure 21a-i**, the green (**Fig. 21a-ii**), red (**Fig. 21a-iii**), and blue (**Fig. 21a-iv**) channel show foci corresponding to the two fork oligonucleotides, the [JF646-Halo-Mcm3]MCMs, and the dCas9-Halo-AF488, respectively. Furthermore, **Figure 21b** shows a kymograph of the foci from **Figure 21a-iii**, showing these bleach in two steps. **Figure 21c-i and c-ii** show the intensity profile corresponding to the left and right foci shown in **Figure 21b**, respectively, showing the red foci bleach in two steps of approximately equal size, indicating the presence of two [JF646-Halo-Mcm3]MCMs and therefore likely a DH at each foci position.

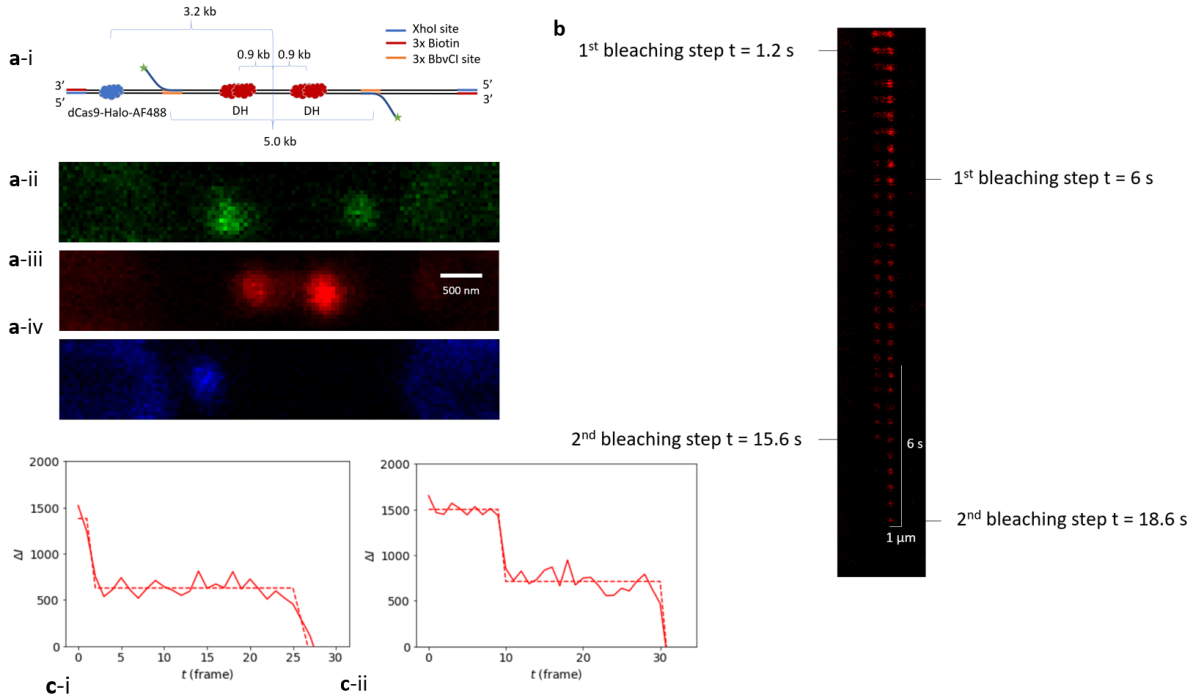


***Figure 21**. (**a**) Averaged intensity per color channel of exemplary construct (**a-i**) Schematic representation. (**a-ii**) Green color channel showing the Cy3 labeled forks. (**a-iii**) Red color channel, showing red fluorescent [JF646-Halo-Mcm3]MCMs (**a-iv**) Blue color channel, showing dCas9-Halo-AF488. Scale bar applies to all fluorescence scans. (**b**) Exemplary kymograph showing the two red foci from **a-iii** bleaching in two steps, indicating presence of two DHs. (**c**) Intensity profiles of foci shown in b, showing DHs bleach in two approximately equally sized steps. (**c-i**) Bleaching trace of left foci in **b**. (**c-ii**) Bleaching trance of right foci in **b**. The dashed lines represent the least squares fit to the bleaching traces. Averaged intensity profiles and the kymograph were obtained using FIJI2.1.0.*

In addition, using the fluorecent fork oligonucleotides as internal nanoruler the distances of the foci corresponding to [JF646-Halo-Mcm3]MCMs and dCas9-Halo-AF488 to the mean center of the constructs was determined and used to compare it to the expectation. From the calculated conversion factor the base pair length was calculated to be (Mean ± SEM) $0.356 \pm 8.2 * 10^{-4}$ nm, which is in good agreement with the value calculated from the previous data set (**see R.3.2.3**).

Moreover, the distances to the center (**Fig. 22**) were calculated to agree well with the expectation. **Figure 22a** shows the positional distributions and a fitted Guassian mixture model of the (Cy3) color channel corresponding to the fluorescent fork oligonucleotides, which was used to find the mean center and base-pair length (**see M.5.2**). The localisation distributions of [JF646-Halo-Mcm3]MCMs and dCas9-Halo-AF488 relative to the mean center are shown in **Figure 22b and c**, respectively. We only fitted a single Gaussian to the localisation data from the dCas9-Halo-AF488, because almost all constructs were trapped in the same orientation and therefore there were very few data points relative to the right of the mean center position.

The mean distances of the [JF646-Halo-Mcm3]MCMs and dCas9-Halo-AF488 differed from the expectations by 14 and 19 nm, respectively (**Table 7**), showing the MCMs form DHs at the HTH sites, as expected. Taken together, this data shows that the fluorescent fork oligonucleotides can be used as internal nanorulers to determine particle positions with sub-pixel precision, the dCas9-Halo-AF488 complex can be used to determine the orientation of single optically trapped constructs, and the data shows DHs can be loaded succesfully at the HTH sites.

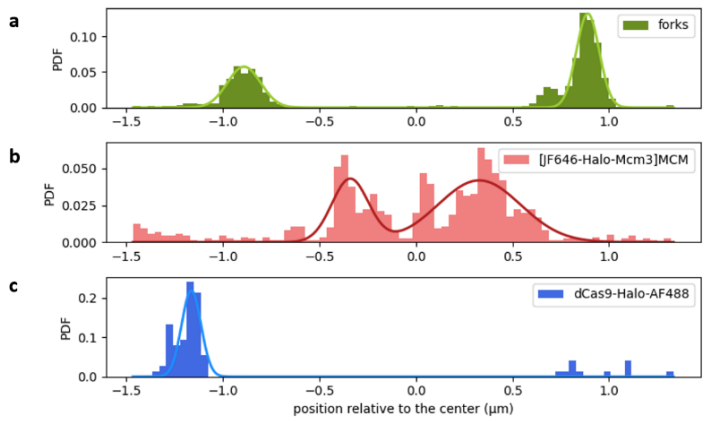


***Figure 22**. Localization distributions of distances to the center from the ensemble of all detected foci over all trapped constructs per color channel, N=13. (**a**) The center between the fitted means of the peaks corresponding to the forks was identified as the mean center of the constructs. (**b**) Distance to the center of red fluorescent [JF646-Halo-Mcm3]MCMs. (**c**) Distance to the center of blue fluorescent dCas9-Halo-AF488.*

Protein	Distance from center (μm)	Distance from center (kb)	Expected distance from center (kb)
[JF646-Halo-Mcm3]MCMs	$0.33 \pm 2.2 * 10^{-2}$	$0.94 \pm 6.2 * 10^{-2}$	0.9
dCas9-Halo-AF488	$1.2 \pm 8.3 * 10^{-3}$	$3.3 \pm 2.3 * 10^{-2}$	3.2

***Table 7**. Calculated (Mean ± SEM) and expected distances from the center of in bulk loaded proteins, N=13.*

R.4 Generation of bacterial overexpression plasmid containing a single-solvent-exposed-cysteine mutant of Pif1

In addition to construct synthesis and characterization, a bacterial overexpression plasmid containing a single-solvent-exposed-cysteine mutant of Pif1 was generated and sequenced. To this end a site directed cysteine to serine mutation was introduced into the previously generated (Daniel Ramírez Montero, personal communication) pET28a-Pif1(40-859) C80S plasmid at amino acid 550 by PCR (**see M.4**). The resulting plasmids were subsequently transformed into NEB® 5-alpha Competent *E. coli* (High Efficiency) cells (NEB Cat. No. 02987). From this transformation five colonies were obtained. The DNA from these colonies was purified using Promega's PureYield™ Plasmid Miniprep System, according to the manufacturers' protocol. The mutagenesis site was then sequenced for each colony to assert whether these took up plasmids with the desired mutation. The results from clone 1 and 2 seemed promising, therefore high-quality DNA was purified from these clones using Marchery-Nagel's NucleoBond Xtra Midi kit, according to the manufacturers' protocol. The entire open reading frame for both the forward and the reverse strands were then sequenced for clone 1, to make sure no other mutations were introduced (**Fig. 23**). The results of the sequencing reactions showed no other mutations were introduced, confirming the bacterial overexpression plasmid containing a single-solvent-exposed-cysteine mutant of Pif1 was successfully produced using clone 1. The plasmids purified from clone 1 were then transformed into chemically competent *E. coli* BL21*DE3 cells which are optimized for protein production.

***Figure 23**. Plasmid map showing the different plasmid features and sequencing results from the forward and backward strand of the entire Pif1 open-reading frame for clone 1. Results are represented by red/yellow arrows, the red regions correspond to successfully sequenced and matching sequencing reads. The figure was made using SnapGene®5.2.4.*

Conclusions

To conclude, the band patterns resulting from the digestions performed on plasmids purified from clones C3 and C4 agree with the expectation and together with the sequencing results show that the desired plasmids are successfully assembled and amplified. The evidence obtained after the biochemical and biophysical characterizations confirms the desired DNA construct was synthesized successfully. Additionally, the densitometry measurements show the forks are incorporated with approximately equal efficiencies at either fork incorporation site.

Furthermore, the findings from the biophysical characterizations support the claim that both forks are incorporated successfully. Increasing, the nicking duration from 3 to 5 h significantly enhanced the fork incorporation efficiency, which was not significantly affected by the increased amount of star activity caused by the Nb.BbvCI nicking enzyme. We also showed that the Cy3 fluorophore takes considerably longer to photo bleach compared to the Cy5 fluorophore, making it more suitable for positional analysis of the forks. It might be interesting to investigate whether the lifetime of the Cy5 fluorophore can be prolonged by using lower laser powers during excitation.

We showed that the MCMs are successfully loaded onto the DNA construct by ORC, Cdc6, and Cdt1 at both artificial head-to-head origin of recognition sites and showed these can form DHs. Loading of the MCMs did not seem to be affected by the forks, neither did the forks seem to be affected by loading of the MCMs. Moreover, the designed gRNA sequences worked well to guide fluorescent dCas9 to a single off-center site and thus allow to ascertain the orientation of single optically trapped construct molecules. It might be interesting to design other gRNA sequences and to further optimize the dCas9 loading procedure to achieve enhanced binding efficiency and reduced off-target binding.

Furthermore, we showed that adding an internal nanoruler, such as the forks, significantly enhances the precision of the positional analysis and allows the determination of particle positions relative to the center of the constructs with sub-pixel (< 50 nm) precision. For future research it may be interesting to determine distances on a single-molecule basis instead of averaging over the ensemble, using the dCas9 complexes to determine the orientation of single optically trapped molecules.

Additionally we conclude that a bacterial overexpression plasmid containing a single-solvent-exposed-cysteine mutant of Pif1 was generated successfully, after sequencing the full open reading frame for both the forward and reverse strands. This mutant version of Pif1 will allow us to site-specifically fluorescently label Pif1 using cysteine-maleimide chemistry, allowing us to study its interaction with the obstructed and converging CMG complexes. The synthesized construct and fluorescently labelled Pif1 could help elucidate whether Pif1 indeed aids the removal of obstructing MCMs.

In the future, it might be interesting to show that CMGs can be loaded onto the construct and ascertain whether converging CMGs are able to unwind the entire construct in the absence of inactive MCMs. Moreover, it would be interesting to directly see CMG translocate over the construct in the presence of MCMs and to see CMG push the DHs towards the center, resulting in obstructed CMG convergence and bypass. Lastly, the constructs could be used to study the movement of CMG combined with other proteins of the replisome and their interaction with the MCMs and Pif1.

References

1. Blumenfeld, B., Ben-Zimra, M. & Simon, I. Perturbations in the replication program contribute to genomic instability in cancer. *Int. J. Mol. Sci.* **18**, (2017).

2. Li, H. & O'Donnell, M. E. The Eukaryotic CMG Helicase at the Replication Fork: Emerging Architecture Reveals an Unexpected Mechanism. *BioEssays* **40**, 1700208 (2018).

3. Deegan, T. D. & Diffley, J. F. X. MCM: One ring to rule them all. *Curr. Opin. Struct. Biol.* **37**, 145–151 (2016).

4. Remus, D. *et al.* Concerted Loading of Mcm2-7 Double Hexamers around DNA during DNA Replication Origin Licensing. *Cell* **139**, 719–730 (2009).

5. Stinchcomb, D. T., Struhl, K. & Davis, R. W. Isolation and characterisation of a yeast chromosomal replicator. *Nature* **282**, 39–43 (1979).

6. Bell, S. P. & Labib, K. Chromosome duplication in Saccharomyces cerevisiae. *Genetics* **203**, 1027–1067 (2016).

7. Méndez, J. & Stillman, B. Perpetuating the double helix: Molecular machines at eukaryotic DNA replication origins. *BioEssays* **25**, 1158–1167 (2003).

8. Douglas, M. E., Ali, F. A., Costa, A. & Diffley, J. F. X. The mechanism of eukaryotic CMG helicase activation. *Nature* **555**, 265–268 (2018).

9. Woodward, A. M. *et al.* Excess Mcm2-7 license dormant origins of replication that can be used under conditions of replicative stress. *J. Cell Biol.* **173**, 673–683 (2006).

10. Courtot, L., Hoffmann, J. S. & Bergoglio, V. The protective role of dormant origins in response to replicative stress. *Int. J. Mol. Sci.* **19**, (2018).

11. Hill, J., Eickhoff, P., Drury, L. S., Costa, A. & Diffley, J. F. X. The eukaryotic replisome requires an additional helicase to disarm dormant replication origins. *bioRxiv* 2020.09.17.301366 (2020).

12. Dewar, J. M., Low, E., Mann, M., Räschle, M. & Walter, J. C. CRL2Lrr1 promotes unloading of the vertebrate replisome from chromatin during replication termination. *Genes Dev.* **31**, 275–290 (2017).

13. Moreno, S. P., Bailey, R., Campion, N., Herron, S. & Gambus, A. Polyubiquitylation drives replisome disassembly at the termination of DNA replication. *Science (80-.).* **346**, 477–481 (2014).

14. Deegan, T. D., Baxter, J., Ortiz Bazán, M. Á., Yeeles, J. T. P. & Labib, K. P. M. Pif1-Family Helicases Support Fork Convergence during DNA Replication Termination in Eukaryotes. *Mol. Cell* **74**, 231-244.e9 (2019).

15. Dewar, J. M., Budzowska, M. & Walter, J. C. The mechanism of DNA replication termination in vertebrates. *Nature* **525**, 345–350 (2015).

16. Low, E., Chistol, G., Zaher, M. S., Kochenova, O. V. & Walter, J. C. The DNA replication fork suppresses CMG unloading from chromatin before termination. *Genes Dev.* **34**,

1534–1545 (2020).

17. Yeeles, J. T. P., Deegan, T. D., Janska, A., Early, A. & Diffley, J. F. X. Regulated eukaryotic DNA replication origin firing with purified proteins. *Nature* **519**, 431–435 (2015).

18. Yao, N. Y. & O'Donnell, M. The replisome. *Cell* **141**, 1088–1088 (2010).

19. Sekedat, M. D. *et al.* GINS motion reveals replication fork progression is remarkably uniform throughout the yeast genome. *Mol. Syst. Biol.* **6**, 1–10 (2010).

20. Echols, H. & Goodman, M. F. REPLICATION. (1991).

21. Sánchez, H. *et al.* DNA replication origins retain mobile licensing proteins. *Nat. Commun.* **12**, (2021).

22. Yardimci, H. & Walter, J. C. Prereplication-complex formation: A molecular double take? *Nat. Struct. Mol. Biol.* **21**, 20–25 (2014).

23. Zhai, Y. *et al.* Open-ringed structure of the Cdt1-Mcm2-7 complex as a precursor of the MCM double hexamer. *Nat. Struct. Mol. Biol.* **24**, 300–308 (2017).

24. Ticau, S. *et al.* Mechanism and timing of Mcm2–7 ring closure during DNA replication origin licensing. *Nat. Struct. Mol. Biol.* **24**, 309–315 (2017).

25. Hua, X. H. & Newport, J. Identification of a preinitiation step in DNA replication that is independent of origin recognition complex and cdc6, but dependent on cdk2. *J. Cell Biol.* **140**, 271–281 (1998).

26. Das, M., Singh, S., Pradhan, S. & Narayan, G. MCM Paradox: Abundance of Eukaryotic Replicative Helicases and Genomic Integrity. *Mol. Biol. Int.* **2014**, 1–11 (2014).

27. Boos, D., Frigola, J. & Diffley, J. F. X. Activation of the replicative DNA helicase: Breaking up is hard to do. *Curr. Opin. Cell Biol.* **24**, 423–430 (2012).

28. Dewar, J. M. & Walter, J. C. Mechanisms of DNA replication termination. *Nat. Rev. Mol. Cell Biol.* **18**, 507–516 (2017).

29. Deegan, T. D., Mukherjee, P. P., Fujisawa, R., Rivera, C. P. & Labib, K. Cmg helicase disassembly is controlled by replication fork dna, replisome components and a ubiquitin threshold. *Elife* **9**, 1–33 (2020).

30. Bouchiat, C. *et al.* Estimating the persistence length of a worm-like chain molecule from force-extension measurements. *Biophys. J.* **76**, 409–413 (1999).

31. Wenner, J. R., Williams, M. C., Rouzina, I. & Bloomfield, V. A. Salt dependence of the elasticity and overstretching transition of single DNA molecules. *Biophys. J.* **82**, 3160–3169 (2002).

32. Baumann, C. G., Smith, S. B., Bloomfield, V. A. & Bustamante, C. Ionic effects on the elasticity of single DNA molecules. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 6185–6190 (1997).

33. Killian, J. L., Ye, F. & Wang, M. D. Optical Tweezers: A Force to Be Reckoned With. *Cell* **175**, 1445–1448 (2018).

34. Heller, I., Hoekstra, T. P., King, G. A., Peterman, E. J. G. & Wuite, G. J. L. Optical tweezers analysis of DNA-protein complexes. *Chem. Rev.* **114**, 3087–3119 (2014).

35. Tinevez, J. Y. *et al.* TrackMate: An open and extensible platform for single-particle tracking. *Methods* **115**, 80–90 (2017).

36. Aaij, C. & Borst, P. The gel electrophoresis of DNA. *BBA Sect. Nucleic Acids Protein Synth.* **269**, 192–200 (1972).

37. Herrero-Galán, E. *et al.* Mechanical identities of RNA and DNA double helices unveiled at the single-molecule level. *J. Am. Chem. Soc.* **135**, 122–131 (2013).

Appendix

Fluorecently labeled fork oligonucleotides

Name	**5'-3' sequence**
DRM 158	[5'Phos]TGAGGTGATTGTAGGCTGAGGCTAGCTCTGCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT[3'Cy5]
DRM 159	[5'Phos]TGAGGTGATTGTAGGCTGAGGCTAGCTCTGCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT[3'Cy3]

Sequencing primers pDRM2; C3 and C4

Sequencing primer	**5'-3' sequence**
DRM 140	AGATTGTACTGAGAGTGCACCATAAAATTG
DRM 142	CCGAACATCGGTCTGATCAACTC
DRM 144	GAGTGAATCTTCTATGGCAAACGAGTC
DRM 177	CGGCCTATTGGTTAACAAATGAGCTG
DRM 180	GAAGGCCACTTCGTAGAAGACCTG
DRM 181	CTGGAACCAGTTCTGGGATTTCC
DRM 182	GCTCATGACAGACAACGGTACC

Mutagenesis primers pET28a-Pif1(40-859)C80S

primer	**5'-3' sequence**
DRM 162	AACATATTTTtctTATGAGGCGC
DRM 163	TCTGGATCCATGAATTCG

Sequencing primers pET28a-Pif1(40-859)C80S/C550S

Sequencing primer	**5'-3' sequence**
DRM 164	GCGAAATTAATACGACTCACTATAGG
DRM 165	CAATATCTCAAGAGAGAAGTTTGG

DRM 166	GAATTCAACTCATCTTCTGTGGG
DRM 167	CCTCCAGAAAAACTCGAGACTTGG
DRM 168	CTAGGATCAAAGCACATCAAAAGG
DRM 169	GTTTAGAGGCCCCAAGGGG
DRM 170	CCACTTTAACTTTTGGAAGTGTCTGACC
DRM 171	GCTATGAAATCGAGTTTATCTAGCAGTTC
DRM 172	CGCTGGTCTTAATGGATTTTTGAACCC
DRM 173	CTAGCTACCGCACTTTCTTCC

crRNA and gRNAs

Name	**5'-3' sequence**
DRM 1	AATTCAATATAAGGCGCGCCATGAAGAGAAGATGGAAGAAGAACTTCATTG
gRNA DRM 1	CACCACGATGCCTGTAGCAATGG
gRNA DRM 2	TACTGAATTCCAAATAGAGGAGG