Characterization of DNA-replication proteins and their molecular mechanisms: a team business

Botto, M.

Chapter 5

Implementation of single molecule light microscopy for the study of multiprotein molecular processes

Margherita Botto[1], John van Noort[2], Meindert H. Lamers[1]

1. Cell and Chemical Biology Department, Leiden University Medical Center, Einthovenweg 20, 2333 ZC Leiden, The Netherlands
2. Biological and Soft Matter Physics, Huygens-Kamerlingh Onnes Laboratory, Leiden University, Niels Bohrweg 2, 2333 CA, The Netherlands

Abstract

To measure the sequential action between different proteins that are part of a molecular process, we need to follow their individual behavior in real time. However, standard biochemical methods measure millions of proteins simultaneously, resulting in the averaging of their behavior and the loss information on their individual actions. In contrast, single molecule light microscopy technologies are ideally suited to follow the behavior of individual proteins. In this work, we describe the implementation of co-localization of single molecule spectroscopy to study the sequential actions of multiple proteins during completion of lagging strand DNA synthesis.

5.1 Introduction

Classical biochemistry methods describe the behaviour of large ensembles of molecules, averaging the measured data over the whole molecular population. However, because the particles involved can have heterogeneous roles and activities, this can lead to a loss of information. In multi steps enzymatic reactions, for example, not all molecules have the same synchronized behaviour. Due to the stochastic nature of molecular interactions, individual events will rapidly become asynchronous. In a bulk experiment the specific behaviours will be averaged, leading to the blurring of the real timeline of enzymatic processes, limiting the discrimination between the different steps involved in the reaction. In recent decades the development of single molecule techniques has helped to understand the details of complicated biological processes, clarifying dynamics and behaviour of individual proteins in multi-protein mechanisms[1–7]. Co-localization of single molecule spectroscopy (CoSMoS) is a single molecule technique that relies on the direct visualization of the interaction between DNA and proteins using fluorescent labeling. By fixing the position of individual DNA molecules on a surface is possible to follow in real time the co-localization between fluorescently labelled proteins and DNA by TIRF microscopy (Fig. 1) and is therefore ideally suited to follow the sequential steps in multi-protein molecular cascades.

One such molecular cascade is Okazaki fragment maturation, a fundamental multi-protein process that happens millions of times in all living organisms during DNA replication. Multiple proteins participate in the Okazaki maturation process, and their individual roles have been characterized[8]. However, what is still missing is how the different steps are temporally organized and how the transition between these different steps occurs. In this work, we describe the implementation of CoSMoS[9] to study the molecular handover between the replicative DNA polymerase Pol I and the ligase LigA on the same DNA molecule. This preliminary work is a vital steppingstone that will allow us to discriminate the different steps present in the Okazaki fragments maturation timeline, showing the full complexity of the mechanism at single-molecule resolution.

5.2 Results

5.2.1 Flow cells

In order to image single molecules in the ~ 100 nm TIRF evanescent field for CoSMoS, we required a small volume sample cell. We therefore created a flow cell by attaching a 25 mm x 75 mm, thickness 1.5H coverslip (ibidi) and a 76 mm x 26 mm x

1.0 - 1.2 mm clear glass microscope slide (Thermo Scientific Menzel-Glaser) using a double-sided tape of 0.12 mm thickness (SecureSeal Adhesive Sheet from Grace Bio-labs) (Fig. 2a).

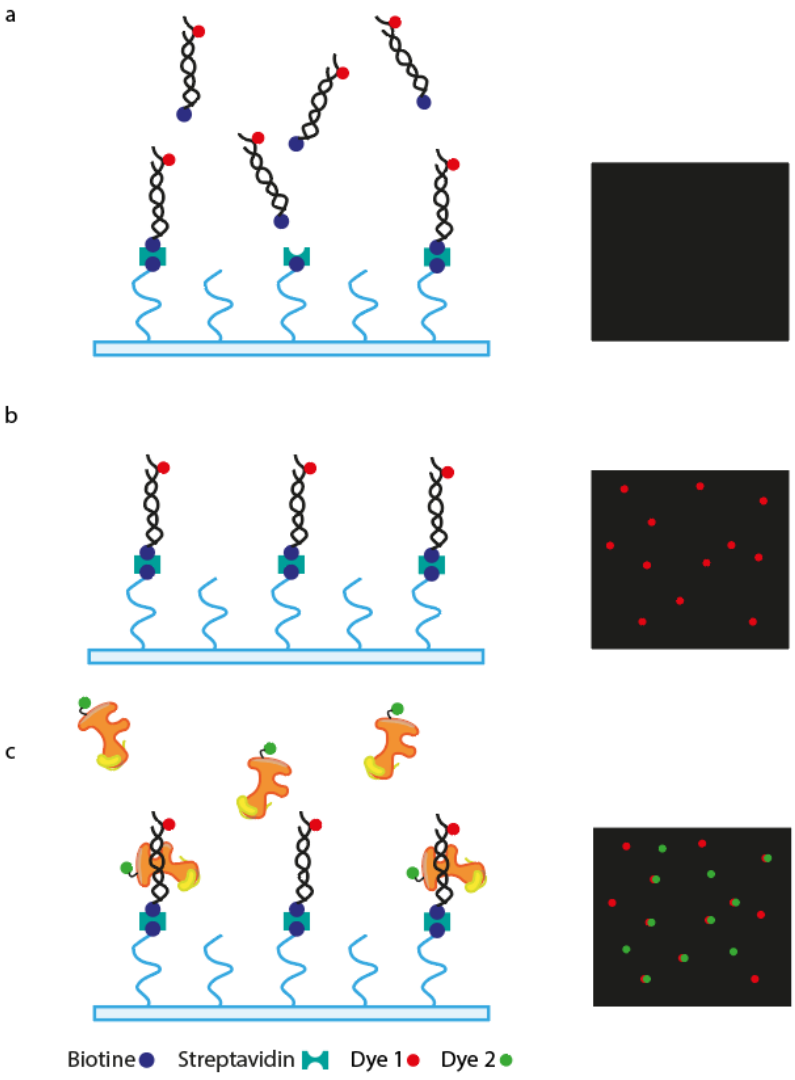

Figure 1. ***Colocalization single molecule spectroscopy. (a)*** *Biotinylated DNA being immobilized on the glass surface because of the biotin-streptavidin interaction.* ***(b)*** *The excess of DNA in solution is being flushed out so that all the DNA visualized is*

immobilized on the surface. ***(c)*** *Labelled protein is being flushed in the imaging chamber. If the protein binds to the DNA we will be able to see the colocalization of the two fluorophores.*

To create channels with a constant volume we cut the double-sided tape with a Silhouette Portrait electronic cutting tool using a pre-designed mold, which helped us cut the tape in a reproducible and fast way. To further improve the reproducibility between experiments, we designed the mold to have three channels so that the same surface-treated coverslip could be used for three different experiments.
To permit the injection of the sample in the imaging chamber before the flow cell assembly we drilled holes of approximately 1.5 mm in diameter in the glass slides at each end of the imaging channels. The holes were big enough to use a pipette tip to add the samples in the flow cell, allowing us to inject new reactions into the flow cell directly on the microscope stage.
During preliminary tests we noticed that often air bubbles were introduced in the imaging channels by pushing the sample in the flow cell, disrupting the surface chemistry and making the channel unusable. To avoid this, we added a plastic cylinder of ~ 0.5 cm in diameter and ~ 1 cm in height on the entrance hole. The plastic cylinder was made by cutting a piece of a 1 ml pipette tip. This way we were able to add the sample in the cylinder at the entrance of the flow cell and pull the reaction mixture out from the other hole with the pipette, letting the reaction pass through the imaging chamber without the risk of creating air bubbles (Fig. 2a-b).
Since directly inserting the pipette into the channel hole proved to be challenging due to lack of space at the microscope, we modified the existing flow cells design in order to be able to connect it to a syringe pump. By gluing an AKTA tube connector (Union 5/16" female - 1/16" male, Cytiva) (see Fig. 2c) directly on the exit channel hole, we found a way to connect the Teflon tubing to the flow cell without the need of epoxy or other glues that could enter into the imaging chamber and make the channel unusable. In this setup, the AKTA tube connector acts as a guide that holds the Teflon tubing in position. The Teflon tubing (1/8" outer diameter) itself is inserted directly into the hole in the glass slide that was drilled to match the width of the tubing. By doing so we firmly fixed the tube to the slide while avoiding the use of any epoxy. Moreover, in our design the tubes are detachable, and once the experiment has ended they can be easily removed and cleaned separately from the rest of the slide.
To be able to create a flow in the imaging chamber, we then connected the free end of the tubing to a 500 µl syringe attached to a KDS single-syringe pump, series 100

(KD Scientific). However, this pump is only capable of pushing the syringe. Therefore, we used a 3D printer (Creality Ender 3) to print adaptors that would fit the syringe on the other side of the pump motor (marked in red in Fig. 2c). This allows us to pull the syringe instead of just pushing it. We thus used the syringe to withdraw the reaction from the reservoir into the flow cell at speeds ~10 µl/sec.

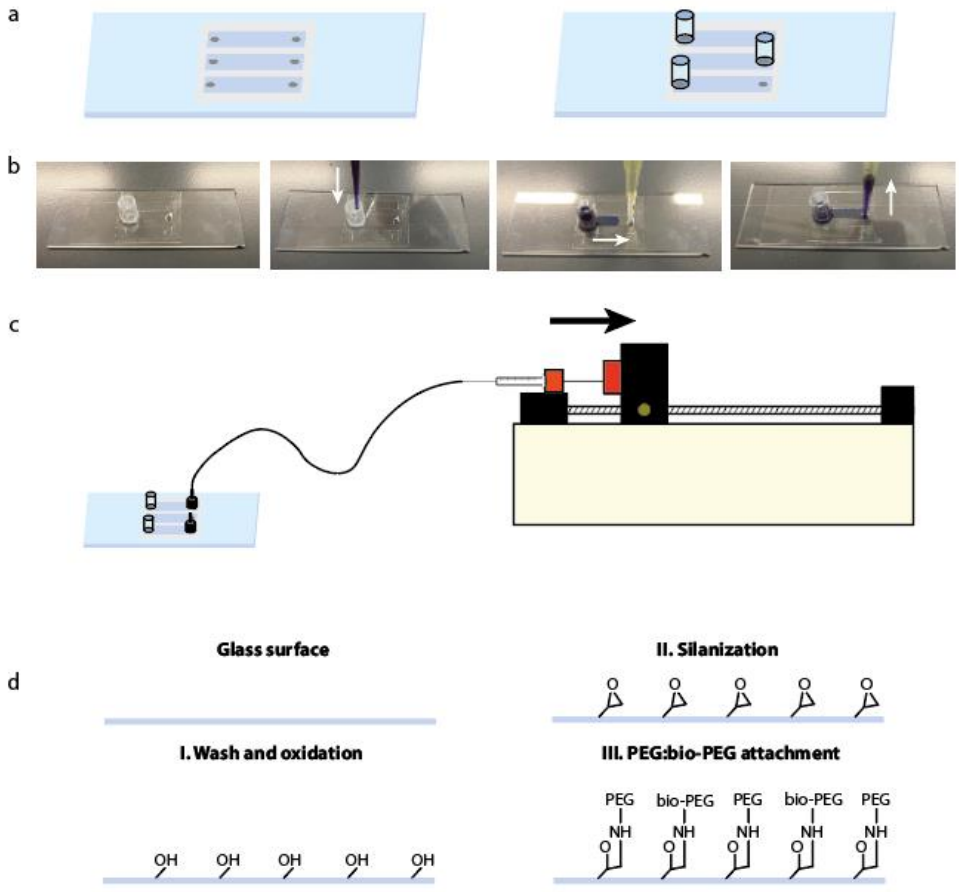

Figure 2. ***Flow cell preparation. (a)*** *Basic flow cell used for our experiment.* ***(b)*** *Injection of the sample into the flow cell by pull with the pipette.* ***(c)*** *Microfluidic flow*

cell set up comprehensive of tubing, syringe and pump. ***(d)*** *Glass surface preparation steps.*

5.2.2 Surface chemistry

Next, to visualize proteins interacting with individual DNA molecules in real time, we had to immobilize the DNA on the glass surface. To do so we first thoroughly cleaned the coverslips using an acetone/methanol mix and Hellmanex (Hellma Analytics), a liquid alkaline concentrate for highly effective cleaning. Then we oxidize the surface using a glow discharger machine (Emitech K950X). Next, to have the correct glass surface chemistry, we silanized the glass using (3-Glycidyloxypropyl)trimethozysilane (GOPTS, Sigma). For protein studies, nonspecific binding to the surface must be suppressed via an additional passivating agent, the most popular being polyethylene glycol (PEG)[10], that can react with the epoxy group present in the GOPTS through a -NH_2 group, forming a covalent linkage. More specifically we have used a mix of PEG (HO-PEG-NH_2) and biotinylated-PEG (CONH-PEG-NH_2) (Rapp Polymere) in a 5:1 ratio, so the DNA-biotin can be attached on the bio-PEG through streptavidin (Fig. 1a). Finally, a layer of bovine serum albumin (BSA, Sigma) and Pluronics (Sigma) was added to further passivate the glass surface. This combination will prevent non-specific hydrophobic interactions of the biomolecules of interest on the glass surface (Fig. 2d).

5.2.3 Fluorophores

A fluorophore is a fluorescent chemical compound that can re-emit light upon light excitation. To image protein-DNA interactions in real time there has to be a way to visually track their position, using of different fluorophores on both of them. Fluorophores with high photostability and long fluorescence lifetime are required in order to visualize protein-DNA interactions for long enough to capture the events of interest. Such fluorophores with will be less prone to photobleaching and enable us to follow binding and dissociation of proteins to DNA over longer time periods (i.e. seconds). The fluorophores used for the experiments present in this chapter are all ATTO dyes as these dyes give well-defined spots that have a significantly long fluorescent lifetime, enough to be ensure that our study was not limited by the photo-bleaching of the fluorophores[11,12].

5.2.4 Protein and DNA labeling

To be able to investigate enzymatic processes in real time, we used site-specific labelling for both DNA and proteins. For proteins labelling we used two different strategies depending on the number of exposed cysteines residues[13]. If the protein of interest contains one exposed cysteine, we labelled it using maleimide-linked dyes that can react almost exclusively with the thiol groups in cysteine residues in the pH range 7.0 - 7.5. Since reducing agents such as dithiothreitol (DTT) contain thiol groups, all DTT was removed by size exclusion chromatography prior to the introduction of the maleimide-dye. Subsequent to the labelling, the protein was re-injected onto a size exclusion column pre-equilibrated in a buffer containing DTT to quench all remaining free dye. This strategy was used to label Pol I klenow fragment (Pol I^{KL}) with a 565 dye, yielding a labelling efficiency of 60 % (Fig. 3a).

Full length Pol I (Pol I^{FL}) contains two cysteines, one in the polymerase domain (Cys 907) and the other in the endonuclease domain (Cys 262). To have a single fluorescent dye on Pol I^{FL}, we mutated the cysteine present on the endonuclease domain. We then labelled it with a 594 ATTO dye and we achieved a labelling efficiency of 70% (Fig. 3c).

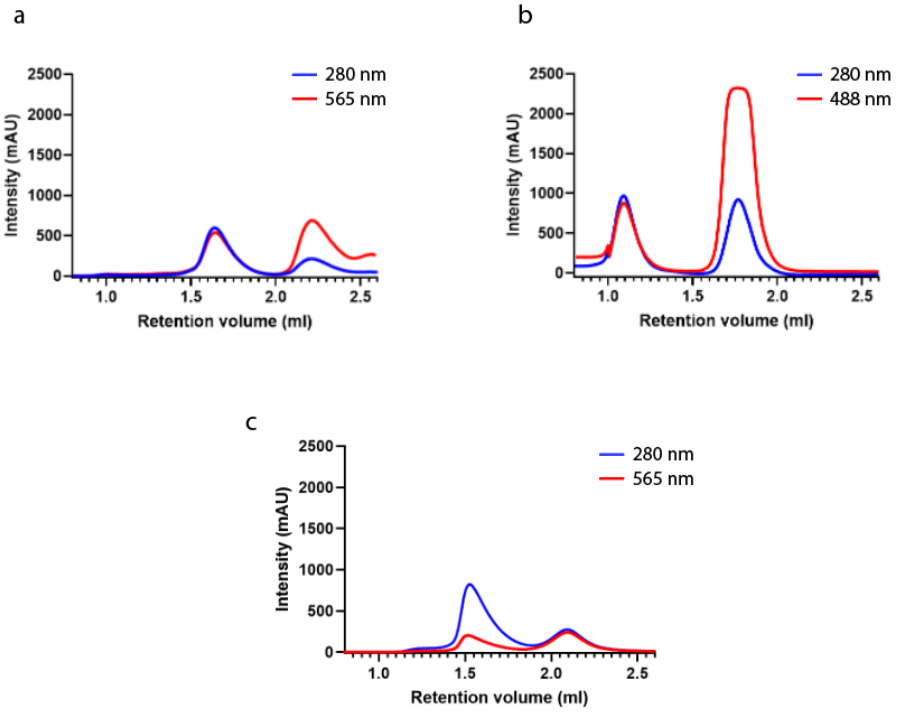

Figure 3. ***Size exclusion chromatography run after labelling of Pol I^{KL} (a) LigA (b) and Pol I^{FL} (c)****. Graph in panel c shows to have a much lower presence of the fluorescent dye compared to the graph in panel a because for Pol I^{KL} it was used the ATTO 565 dye, while for Pol I^{FL} the ATTO 595, that gets excited at a slightly higher wavelength.*

For proteins containing more than two surface-exposed cysteine residues we used another labelling strategy. Removing all the cysteines residues would have meant multiple mutagenesis rounds, and risking to affect the protein functionality. Therefore we added a Ybbr sequence[14] to the N-terminus of the protein. This sequence was identified from a genomic library of *Bacillus subtilis* by phage display as an efficient substrate for Sfp phosphopantetheinyl transferase-catalysed protein labelling by small molecule-CoA conjugates. Therefore, when a maleimide-functionalized dye is covalently linked to the free sulphur in CoA, the resulting CoA-dye complex can be enzymatically conjugated to the protein of interest. We used this strategy to label LigA with a 488 dye and we achieved a labelling efficiency of 70 % (Fig. 3b).

After fluorescent labelling of the proteins we performed control experiments over their activity (Fig. 4).

The DNA substrates were labelled using an amino-modified thymine and NHS ester-activated dyes that react with the primary amines. We labelled our ssDNA template with a 647N ATTO dye and achieved a DNA labelling efficiency of 81% (Fig. 5).

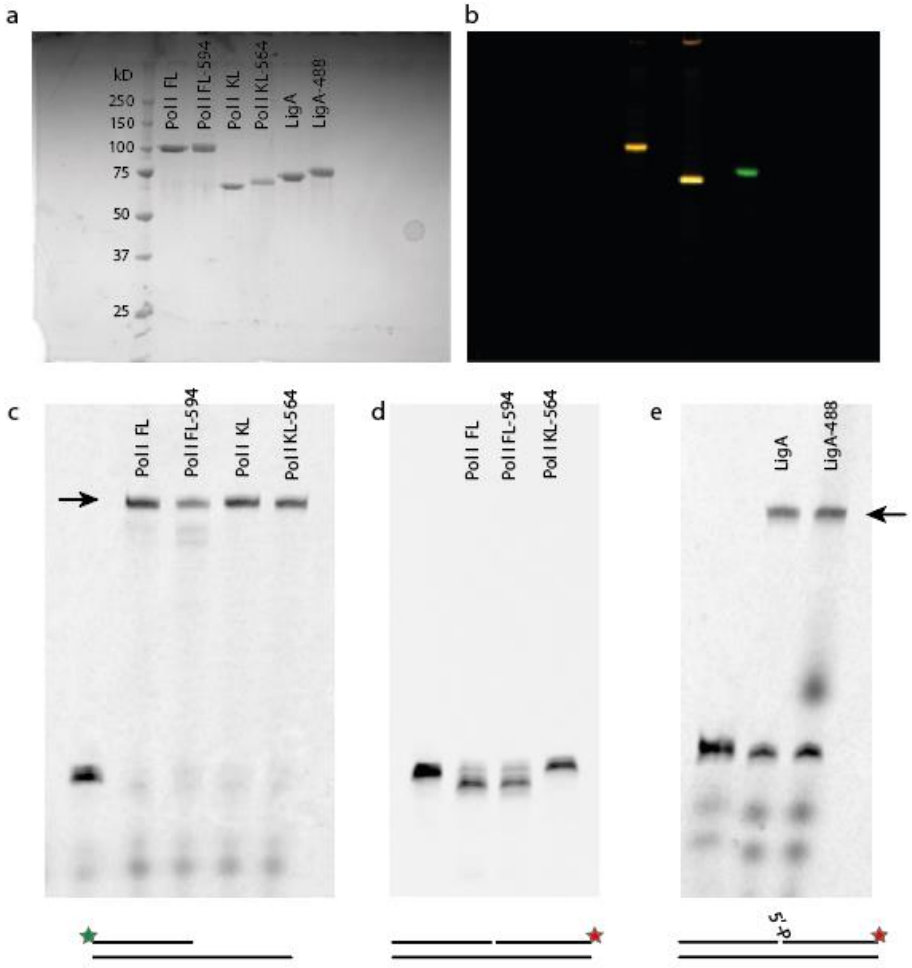

Figure 4. ***The fluorescently labelled proteins retain wild-type activity.*** ***(a)*** *SDS-PAGE gel stained with Coomassie blue showing the labeled and unlabeled proteins.* ***(b)*** *Gel in panel a scanned with 555 wavelength (for Pol I^{FL} and Pol I^{KL}) and with 488 (for LigA) before Coomassie staining.* ***(c)*** *Primer extension assay showing that the labeled Pol I^{FL} and Pol I^{KL} retain the same activity as the unlabeled ones. Arrow indicates final products.* ***(d)*** *Incubation of a DNA substrate in absence of dNTP's to show that the endonuclease of the labeled Pol I^{FL} retain wild-type activity. Labeled Pol I^{KL} doesn't show any endonuclease activity, as expected.* ***(e)*** *Ligase activity assay to show that the labeled LigA retain the same activity as the unlabeled one. Substrates used for experiments in panel c-e can be found in Supplementary Table 2.*

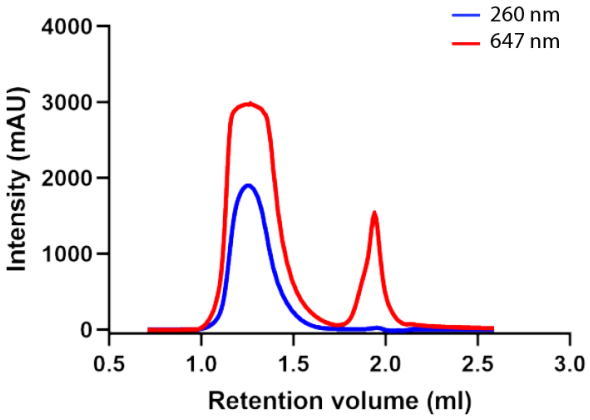

Figure 5. *Size exclusion chromatography after labelling of ssDNA with ATTO647N.*

5.2.5 Data acquisition

The labelled proteins where injected into a flow cell using a syringe pump and data acquisition was started immediately. An initial image stack was collected using two laser channels, one for the DNA and the other for the labelled protein that serves to align the two colour channels and map the position of the DNA molecules. The two channels were acquired alternatingly on the same Zyla sCMOS camera and a 50 ms exposure for channel. Because of the mechanical movement of the filter wheel, there is a delay of 150 ms per channel, leading to a 200 ms time per channel and 400 ms per frame in total when two channels were in use (Fig. 6a).

Following the initial image stack, a second image stack was recorded in the same field of view using only the protein channel in order to avoid missing out on protein binding events that are faster than the 400 ms frame rate. Using a single colour channel, each frame was now 50 ms long and we imaged 500 frames each field of view (Fig. 6b).

5.2.6 Data analysis

The analysis of multi-color image stacks involves multiple steps, and a specific analysis method was developed in collaboration with John van Noort and it is briefly described here.

As discussed above, for each field of view two image stacks were recorded: one using both protein and DNA channel followed by an image stack using just the

protein channel (Fig. 6a-b). For each field of view acquired with the dual color measurement, the different laser channels were aligned and corrected for drifting of the microscope stage in order to be sure that the DNA spot acquired for the 637 nm channel corresponded to the same spot in the protein channel. The position of each DNA molecule was then mapped and stored (Fig. 7a). Mapping of DNA positions was done by iteratively scanning each field of view for the highest intensity peak, record its coordinates, set it to zero and look for the next highest intensity peak throughout the entire field of view. This was performed until a lower threshold was reached where peaks could no longer be distinguished from the background noise.

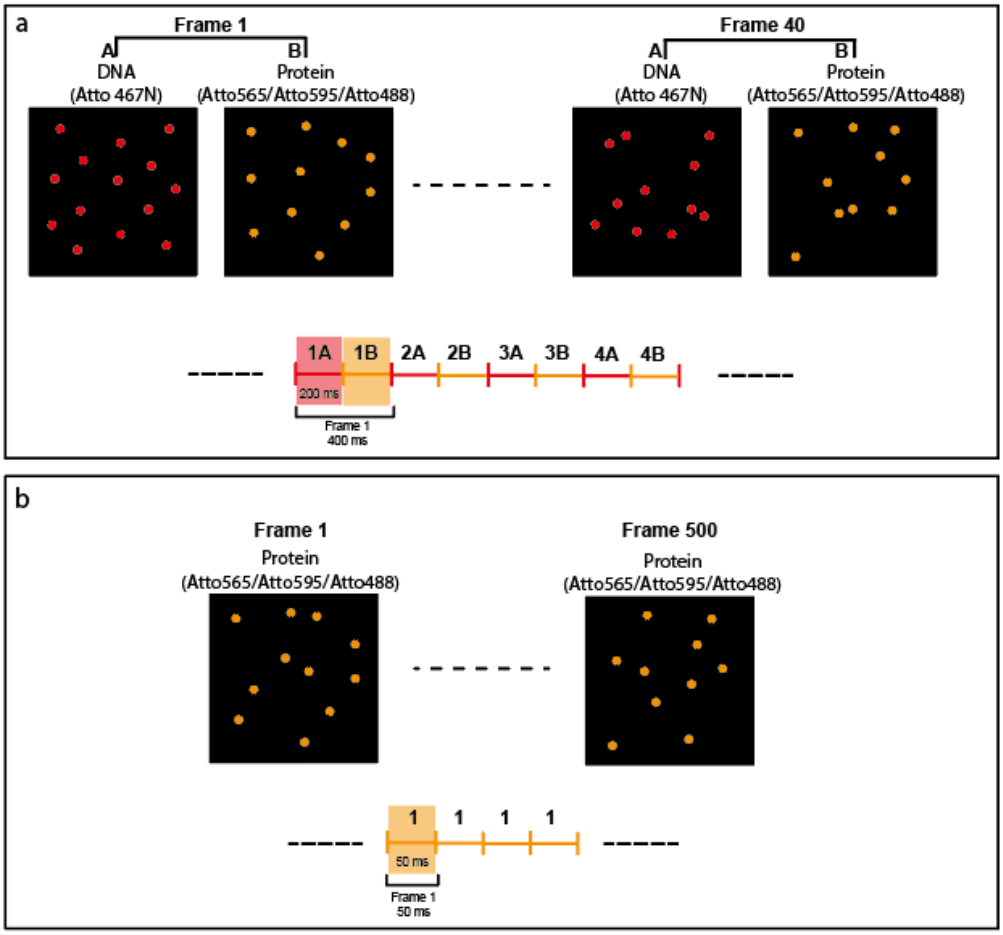

Figure 6. ***Data acquisition scheme.*** ***(a)*** *An initial image stack of the DNA and the protein channel is acquired sequentially. The time between two consecutive frames of the laser channel when two channels are used is 400 ms.* ***(b)*** *A second image stack is acquired for the same field of view using only the protein channel for faster time resolution. Data acquired over 500 frames.*

Next, for each field of view we analyzed the single color channel by using the DNA positions from the two-channel measurement as a reference. After drift correction, time traces were extracted at each DNA position and examined for protein colocalization events. For this, each time trace was fitted using a two states hidden-Markov model[15,16](HMM), which provides a method to analyze time series data. For each time trace recorded on a DNA position, we used the HMM fitting to extract each instance where a higher signal was observed, reflecting a protein binding event. The binding time was then stored as the lifetime τ (Fig. 8b).

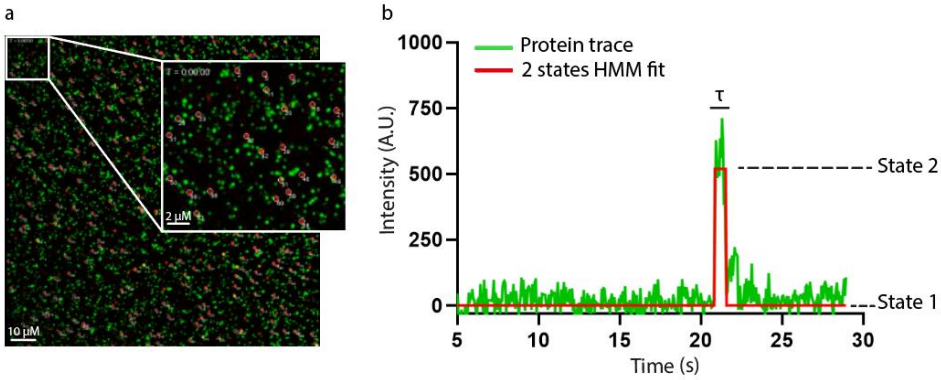

Figure 7. ***Data analysis strategy.*** **(*a*)** *Example of a field of view acquired using the two-channels method. The DNA molecules are represented by red dots and the protein by green dots.* ***(b)*** *Example of protein trace (in green) acquired over a DNA molecule position, fitted using the 2-state Hidden Markov model (HMM) (in red).*

5.2.7 Preliminary CoSMoS data acquisition

As a preliminary experiment, we started by testing a doubly labelled DNA on a TIRF microscope. We labelled the DNA with both ATTO 647N and ATTO 488 because of their long lifetime and stability (Fig. 8a and substrate T647-P488 in Supplementary

Table 1). We then imaged each color channel for 50 ms. Data in Fig. 8b-d shows that not all the dots were colocalizing within the two different channels. This can be explained by multiple reasons. First it is possible that the annealing of the two strands was not exactly at a 1:1 ratio, therefore the difference in number of fluorescent dots between the channels. Secondly there is the possibility of a suboptimal labelling of one of the two strands, leading to the observation of just one DNA strand instead of two. Nevertheless, when we measured the amount of fluorophores that colocalize in various fields of view, on average we have 60% of dots colocalizing in the two channels. Therefore, we can conclude that we are able to visualize colocalization between the two channels.

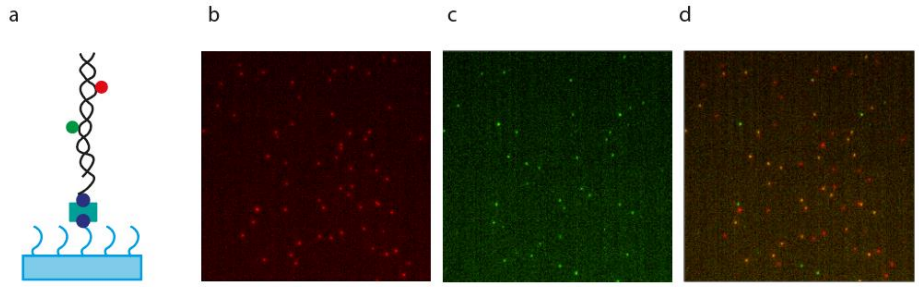

Figure 8. ***Colocalization measured on a double labelled dsDNA. (a)*** *Schematic representation of the double stranded substrate used for the experiment.* ***(b & c)*** *Signal given by ATTO 647N and ATTO 488.* ***(d)*** *Signal given by the merged channels. The orange reflects the dots that show colocalization.*

5.2.8 Protein-DNA colocalization

Next we wanted to test if our flow cell preparation, data collection and data analysis allows for the detection of the activity of a single protein on a single DNA molecule in real time. For this we have studied the behavior of *E. coli* polymerase Pol I and LigA on DNA. It is known that Pol I has a great affinity for DNA[17–19]. Therefore, we started by testing Pol I^{FL} interaction with a nicked DNA substrate in real time (substrate Nicked T647 in Supplementary Table 1). Out of 308 colocalization events we determined that the lifetime (τ) of Pol I^{FL} on this DNA substrate is 1 second

Next, we studied the behavior of LigA on the same nicked DNA substrate. Previous studies reported a K_D of *E. coli* LigA for a nicked DNA of ~5 μM[20], indicating a low

affinity for the substrate. Fig. 9d shows for LigA a lifetime of 0.3 seconds, confirming the lower affinity of LigA for a nicked DNA substrate compared to Pol I^{FL}.
Finally, we investigated the exchange between Pol I^{FL} and LigA on the same DNA molecule. To do this we have used a labelled Pol I^{FL} together with an unlabeled LigA, in order to visualize how the behavior of Pol I^{FL} can be affected by the presence of LigA in real time. Fig. 10b shows that the lifetime of Pol I^{FL} on a nicked DNA is reduced in presence of LigA. Hence in absence of dNTP's and on a nicked DNA substrate, the two proteins appear to compete for the same DNA molecule, leading to a shortening of the lifetime of Pol I^{FL} on DNA.

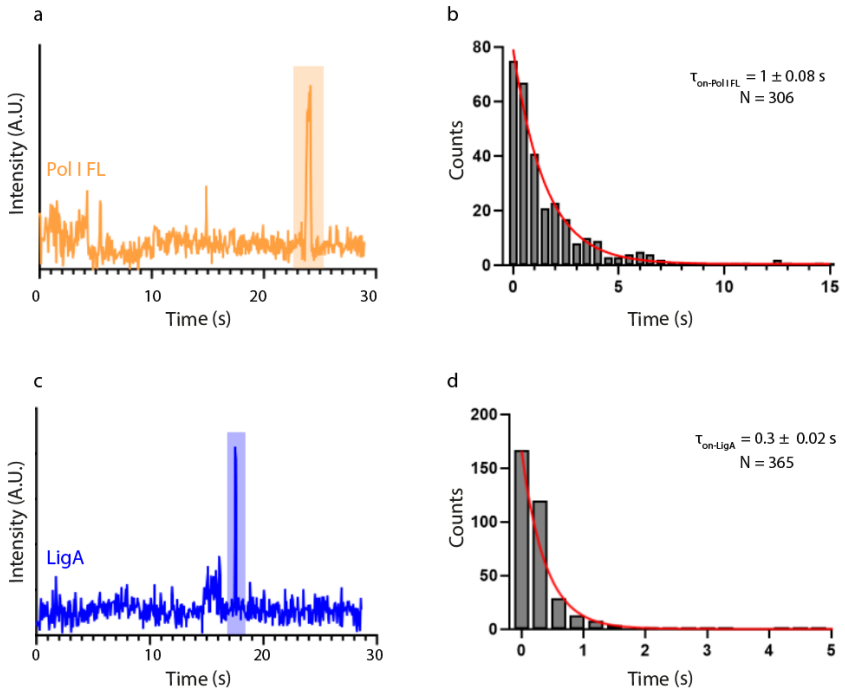

Figure 9. ***Pol I^{FL} and LigA binding on a nicked DNA. (a)*** *Graph showing the signal intensity of Pol I^{FL} at the position of a DNA molecule.* ***(b)*** *Histogram of the lifetime of 308 binding events of Pol I^{FL} on a nicked DNA.* ***(c)*** *Graph showing the signal intensity of LigA at the position of a DNA molecule.* ***(d)*** *Histogram of the lifetime of 365 binding events of LigA on a nicked DNA.*

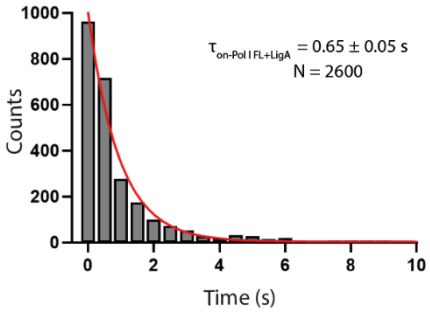

Figure 10. ***Pol I^{FL} + LigA binding on a nicked DNA***. *Histogram of the fluorescent lifetime of Pol I^{FL} on a nicked DNA when LigA is also present.*

5.3 Future implementations

Throughout our measurements we observed the presence of non-specific protein binding to the glass slide surface. The fact that all the proteins tested showed the same behavior means that the non-specific binding is not caused by individual protein properties but instead by the glass slide preparation or the buffer used. However, the non-specific binding did not affect our measurements because of our data acquisition method. By measuring each field of view twice, first with a short two color measurement and then with a long one-laser measurement, we first bleach all the proteins bound to the surface preparing the field of view for the long measurement, where we only visualize the proteins binding and releasing the DNA molecules. Even if the non-specific binding of the proteins on the glass surface was not relevant for our data acquisition and analysis, it would be beneficial for future experiments to reduce the non-specific signal, especially when we are looking rare events, or look at early stages of a molecular reaction where bleaching of the non-specific signal is not an option. Higher PEG densities isolate the biomolecules from the surface more efficiently, so it is possible that the PEG amount used in this chapter for slide preparation didn't create a dense enough barrier to avoid the sticking of the proteins on the glass surface. Because of this, a range of different PEG ratios should be tested. Other interesting conditions to test could be a variation of the concentration of the blocking reagent BSA or a pH buffer variation. Alternatively, it has been reported that it is possible to remove the non-specific signal during the

data analysis. Previous studies[21,22] have showed that if the behavior of nonspecific and specific binding molecules is different enough, during data analysis it would be possible to discriminate between the two populations and thus remove the unwanted data.

When we compare our results with the ones reported in [23], the lifetime measured here for Pol I^{FL} is ~40 times lower. One possible explanation for this difference could be the position of the fluorophore on the DNA. In the substrate used for the above experiments the dye is four nucleotides away from the nick, close to the binding site of the polymerase. This could potentially disrupt the polymerase binding on DNA and lower its lifetime. For this reason, we tried to move the dye 10 nucleotides from the nick. However, the lifetime of Pol I^{FL} did not change (data not shown). To solve this issue additional experiments should be performed.

The microscope used for the above experiments also has the possibility to use two cameras to measure two different channels simultaneously, thus decreasing the acquisition time per frame. Moreover, since every camera has its own filter wheel, the further time delay per channel due to its mechanical rotation would be eliminated. Unfortunately, in our experiments this was not possible because of the combination of dyes. The image splitter between the two cameras has a cutoff at 560 wavelength and this makes impossible to measure Pol I (labelled with ATTO 565 or ATTO 595) and the ATTO 647N-DNA at the same time on the two different cameras. Therefore, in future experiments a dye with a lower emission wavelength could be used to speed up the two-color data collection.

In this chapter we showed that our CoSMoS setup allowed us to visualize, track and analyze labeled proteins at a single-molecule resolution and in real time. However, information about how the proteins involved in Okazaki fragment maturation coordinate their activities in real time is still lacking. With the technology described in this chapter we are now able to start analyzing the time dependent activities of Pol IIIα, Pol I and LigA on the same DNA molecule. If each element is labeled with a different dye, this could be performed by the use of three different laser channels: one for the DNA, one for Pol IIIα and one for Pol I. Therefore, it would be of great value to also implement a three color-channel data analysis.

Finally, our analysis doesn't allow an automatic discrimination between single and multiple protein DNA binding events. This can be easily done by looking at the intensity of the signal from the protein channel. In our experiments when two or more proteins would bind to the same DNA molecule the data would show a clear doubling or tripling of the signal detected for a single binging event. An automatic implementation for this matter would be a helpful tool.

5.4 Discussions

In this chapter we have implemented and developed methods for single molecule flow cell preparation, for data acquisition and data analysis. While multiple slides surface preparation methods have been developed in the past[24–28], we have modified these so that they allow an easy and reliable slides preparation pipeline. Moreover, to adjust to the space requirements of the microscope, we designed new easy-to-build flow cells that give the possibility to measure over two different channels using a syringe pump creating a constant flow in each imaging chamber. Lastly, because of the lack of commercially available data analysis programs, our own data acquisition and data analysis method has been developed in collaboration with John van Noort. Because of this, we were able to study the behavior of Pol I and LigA on a nicked DNA at single molecule resolution. The data acquired showed that Pol I^{FL} has a longer lifetime on a nicked DNA compared to LigA in absence of dNTP's. Next, we measured that the presence of LigA influences the behavior of Pol I on DNA, suggesting that LigA competes with Pol I^{FL} for the binding of a nicked DNA substrate.

In conclusion thanks to the developments presented in this chapter, it will be possible to perform single molecule experiments over a large variety of proteins interacting with DNA. This technology will be valuable to several other projects in our laboratory that study multi-protein molecular pathways, such as DNA mismatch repair, the mutasomal complex from *Mycobacterium tuberculosis*, and finally, the actions of Pol IIIα, Pol I, and LigA during Okazaki fragment maturation.

5.5 Materials and Methods

5.5.1 Chemicals and reagents

All chemicals were purchased from Sigma, unless indicated otherwise. DNA oligonucleotides were purchased from IDT. Chromatography columns were purchased from Cytiva.

5.5.2 Cloning

To label Pol I^{FL}, the cysteine on the endonuclease domain was removed with the use of mutagenesis (see Supplementary Table 3).

For labelling purposes the Ybbr sequence (DSLEFIASKLA)[13] was added on the X-terminal to LigA in a pETNKI-his3C-LIC-amp vector (see Supplementary Table 3).

5.5.3 Proteins expression and purification

All proteins were expressed in *E. coli* BL21 (DE3) (Novagen) for two hours at 37 °C. All cell pellets were lysed by sonication and clarified by centrifugation at 24,000 x g. Protein purification was performed using buffer A (25 mM Imidazole pH 7.5, 500 mM NaCl, 2mM DTT, 5% Glycerol), buffer B (25mM Tris pH 8.0, 2 mM DTT, 10% Glycerol), buffer C (HEPES pH 7.5, 100 mM NaCl, 2mM DTT), and buffer D (25mM Tris pH 8.5, 2 mM DTT, 10% glycerol), with addition of imidazole or NaCl as indicated. All purified proteins were flash frozen in liquid nitrogen and stored at -80 °C.

Pol I^{FL} & Pol I^{KL}, were purified using a HisTrap column pre-equilibrated in buffer A and eluted using a gradient to 500 mM Imidazole in buffer A. The his-tag was removed by overnight digestion, buffer exchanged to buffer A (at 25 mM Imidazole) and followed by a second Histrap column to remove undigested protein. The flowthrough was injected onto a Hitrap Q column pre-equilibrated in buffer B with 150 mM NaCl and eluted with a gradient to 1 M NaCl in buffer B. Pol IIIα was further purified using a Superdex 200 size exclusion column pre-equilibrated in buffer C.

LigA was purified using a HisTrap column pre-equilibrated in buffer A and eluted using a gradient to 500 mM Imidazole in buffer A. The eluted protein was diluted 5-fold to 100 mM NaCl in buffer B, injected onto a Hitrap Heparin column pre-equilibrated in buffer B with 100 mM NaCl and eluted with a gradient to 1 M NaCl in buffer B.

5.5.4 Protein labelling

LigA was labelled at a single site in the N-terminus using the Sfp method. ATTO 488 fluorophore was conjugated to Coenzyme A[29–31]. Next, LigA was labelled in a reaction containing 10 mM MgCl2, 50 mM Hepes 7.5, 10 mM DTT, 2 µM Sfp, 100 µM CoA-ATTO 488 and 50 µM protein. The CoA-dye-LigA reaction was incubated for 3 hours at room temperature. The labelled protein was purified from the free dye by gel filtration using a 2.4 ml Superdex 75 column (Cytiva). The labelling efficiency was determined spectrophotometrically using the protein and the fluorophore absorption ratios. The labeling efficiency of LigA was measured to be 70 %.

Pol I has two surface-exposed cysteine residues, one in the polymerase domain and one in the endonuclease domain. For Pol I^{FL} labeling we removed the cysteine in the endonuclease domain by mutagenesis. Then Pol I^{FL} was labeled using ATTO 594-maleimide dye.

Pol I^{KL} was instead labeled using ATTO 565-maleimide dye. To further proceed with the labeling procedure we buffer exchanged Pol I^{FL} and Pol I^{KL} by gel filtration on a 2.4 ml Superdex200 Increase column (Cytvia) to remove DTT. The proteins were

then labelled using a 5 molar excess of dyes for 2 hours at room temperature in 20 mM Hepes 7.5 and 50 mM KGlu. The labelled protein was then purified away from the free dye by gel filtration using a 2.4 ml Superdex 200 column (Cytiva) in buffer containing DTT to quench the labelling reaction. The labelling efficiency for Pol I^{FL} was 70 % and for Pol I^{KL} was 60 % and was measured spectrophotometrically.

5.5.5 Single molecule DNA substrates

The DNA oligos were purchased from IDT. The template used for the single molecule experiments above is biotinylated at 5' end to stick to the surface and presents an internal amino modification on a thymine base (amino dT) for the labelling of the DNA. It was then labelled with an excess NHS-ester activated ATTO fluorophores for 1 hour at room temperature in 0.1 NaH_2PO4 and 0.1 M Na_2HPO4 pH 7. The free dye was separated by gel filtration on a 2.4 ml Superdex 75 column (Cytiva). Finally, the labelled template DNA was then annealed in 10 mM Tris pH 7.5, 50 mM NaCl and 1 mM EDTA to a primer DNA in a 1:1 ratio (Substrates in Supplementary Table 1).

5.5.6 Primer extension assays

Primer extension assays were performed using oligonucleotides shown in Supplementary Table 2. All assays were performed in 20 mM HEPES pH 7.5, 2 mM DTT, 5 mM $MgCl_2$, 50 mM NaCl and 0.05 mg/ml BSA. For gel analysis, reactions were performed at 22 °C with 100 nM protein and 50 nM of DNA substrate. Primer extensions were carried out for 1:30 minutes in the presence of 100 µM dNTP's (otherwise stated). Reactions were stopped in 35 mM EDTA and 65 % formamide and separated on a denaturing 20% acrylamide/bis-acrylamide (19:1) gel with 7.5 M Urea in 1x TBE for 80 minutes at 30 W. The gel was imaged with a Typhoon Imager (GE Healthcare).

5.5.7 Ligase activity assay

Ligase activity assay was performed using oligonucleotide shown in Supplementary Table 2. The assay was performed in 20 mM HEPES pH 7.5, 2 mM DTT, 5 mM $MgCl_2$, 50 mM NaCl and 0.05 mg/ml BSA. To start the reaction 300 nM of LigA was added to 100 nM DNA mixed with 50 µM NAD^+. The reactions were stopped after 45 minutes at RT in 35 mM EDTA and 65 % formamide and separated on a denaturing 20% acrylamide/bis-acrylamide (19:1) gel with 7.5 M Urea in 1x TBE for 80 minutes at 30 W. The gel was imaged with a Typhoon Imager (GE Healthcare).

5.5.8 Flow cells preparation

Glass slides and coverslips were washed and sonicated in water for 20 minutes. Next the coverslips were sonicated in an acetone/methanol mix (70%/30% v/v) for 20 minutes. After washing and sonicating the coverslips several times with water to remove the acetone/methanol, they were sonicated for 20 mins in milliQ + 4% Hellmanex (Hellma Analytics). Next the coverslips were carefully washed in water to remove the Hellmanex and sonicated in milliQ for 10 mins. To oxidize the coverslips surface we placed the slides in an air glow discharger machine for 60 seconds at 20 mA. Next, we placed the coverslips on a Petri dish on a heating plate at 100 °C, paying attention to put looking upwards the face that has been glow discharged. To silanize the glass surface we pipetted on each slide 10 µl of GOPTS in the center of each coverslip and we placed another coverslip on top of each one to create a sandwich. Immediately after finishing to prepare the sandwich we placed the Petri dish into a 75 °C oven for 20 minutes. Next, we placed the Petri dish back on the heating plate at 100 °C and we carefully separated the two coverslips and quickly put them in a holder filled with acetone to avoid exposure to air as much as possible. While submerged in acetone, we moved the slides up and down to remove the excess of GOPTS. In the meantime we prepared the PEG solution by dissolving a mix of PEG and Bio-PEG in a 1:5 ratio in 100 mM Sodium Bicarbonate. Finally, we dry the coverslips and put them in a humid container with the prepared face looking up. We then add 100 ul of PEG solution on each slide and lay another slide on top of each to create a sandwich. We leave them overnight in the humid container with no exposure to light and the day after we open the sandwiches and put the slides in water. We removed the excess of PEG solution by moving the slides up and down and finally we dried them using a nitrogen flow.

The imaging chamber (15 µl) was assembled by creating a sandwich between the prepared coverslip and a glass slide using a double face tape shaped with three channels. The channels were then flushed with 30 µl of 20 mM Tris pH 7.5, 50 mM potassium glutamate, 8 mM $MgCl_2$, 4% Glycerol, 2 mM DTT, 0.01% Tween 20 and 1 mM Trolox (Imaging Buffer). The channels were further passivated injecting 10 mg/ml BSA (New England Biolabs) and 1% Pluronics (Sigma). Streptavidin (0.2 mg/ml) (New England Biolabs) was added last to bind the biotin-DNA. Finally, each channel was flushed with 30 µl of Imaging buffer to remove the excess of BSA, Pluronics and Streptavidin.

5.5.9 TIRF microscopy

The DNA was injected into the imaging chamber using a KDS single-syringe pump, series 100 (KD Scientific) using a flow rate of 500 µl/min.

Each flow cell was imaged on an Andor Dragonfly 500 microscope (Oxford Instruments) using a HC PL APO 100x/1.47 OIL CORR TIRF objective (Leica). The laser used were 488, 561 and 637. Movies were acquired on a Zyla sCMOS 4.2 2048 x 2048 camera with an exposure of 50 ms per frame with rapid alternation between the laser channels (150 ms). The frame rate was 400 ms when two channels were in use.

5.5.10 CoSMoS experiments

All the CoSMoS experiments were performed in 300 pM DNA and 250 pM proteins. The experiment performed using labelled Pol I^{FL} and unlabeled LigA was performed by mixing the two proteins at a final concentration of 250 pM before the injection and then injected together to start the experiment.

Supplementary material

Substrate name	Length	Sequence
T647-P488	33 27	5'Bio- CATAATATCCATGCTTCACCTTCATCCAAATCC 3' 3'488-TAGGTACGAAGTGGAAGTAGGTTTAGG 5'
Nicked T647	33 17 + 16	5'Bio- CATAATATCCATGCTTCACCTTCATCCAAATCC 3' 3' GTATTATAGGTACGAAGTGGAAGTAGGTTTAGG 5'

Supplementary Table 1. *Primers used for CoSMoS experiments. In red is highlighted the amino modification at the thymine base used to label the primer using NHS-ester-activated Atto-647N dye. Underlined is the position of the nick.*

Substrate name	**Length**	**Sequence**
Cy3_Extension	54 24	5'CGCTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACC 3' 3'CTGACCCTTTTGGGACCGCAATGG Cy3-5'
Cy5_Endonuclease	54 30 + 24	5'CGCTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACC 3' 3'-Cy5 GCGAGTGACCGGCAGCAAAATGTTGCAGCACTGACCCTTTTGGGACCGCAATGG 5'
Cy5_Ligase	54 30 + 24	5'CGCTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACC 3' 3'-Cy5 GCGAGTGACCGGCAGCAAAATGTTGCAGCApCTGACCCTTTTGGGACCGCAATGG 5'

Supplementary Table 2. *Primers used for primer extension assay and Ligase activity assay. Underlined is the position of the nick. Phosphorylation at 5' of the Cy5-labelled primer indicated with a p.*

Substrate name	Length	5’ to 3’ Sequence
Pol1_NoCys_ FW	27	CTGGAGCTGACCGCTGAACAACTGGAA
Pol1_NoCys_ RV	27	TTCCAGTTGTTCAGCGGTCAGCTCCAG
LigAYBBR_FW	41	TTTATTGCTAGTAAACTGGCGATGGAATCAATCGAACAACA
LigAYBBR_RV	41	AGCAATAAATTCAAGAGAATCCCCGGGTCCCTGAAAGAGCA

Supplementary Table 3. *Primers used for mutagenesis.*

References

1. Xue, Q. Differences in the chemical reactivity of individual molecules of an enzyme. *Nature* **373**, 681–683 (1995).
2. Craig, D. B., Arriaga, E. A., Wong, J. C. Y., Lu, H. & Dovichi, N. J. Studies on single alkaline phosphatase molecules: Reaction rate and activation energy of a reaction catalyzed by a single molecule and the effect of thermal denaturation - The death of an enzyme. *J. Am. Chem. Soc.* **118**, 5245–5253 (1996).
3. Edman, L., Mets, Ü. & Rigler, R. Conformational transitions monitored for single molecules in solution. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 6710–6715 (1996).
4. Ha, T., Enderle, T., Chemla, D. S., Selvin, P. R. & Weiss, S. Single molecule dynamics studied by polarization modulation. *Phys. Rev. Lett.* **77**, 3979–3982 (1996).
5. Macklin, J. J., Trautman, J. K., Harris, T. D. & Brus, L. E. Imaging and time-resolved spectroscopy of single molecules at an interface. *Science (80-.).* **272**, 255–258 (1996).
6. Weiss, S. Fluorescence Spectroscopy of Single Biomolecules. *Science (80-.).* **286**, 2098–2102 (1999).
7. Biju, V., Micic, M., Hu, D. & Lu, H. P. Intermittent single-molecule interfacial electron transfer dynamics. *J. Am. Chem. Soc.* **126**, 9374–9381 (2004).
8. Lewis, J. S., Jergic, S. & Dixon, N. E. *The E. coli DNA Replication Fork*. *Enzymes* **39**, (Elsevier Inc., 2016).
9. Friedman, L. J., Chung, J. & Gelles, J. Viewing dynamic assembly of molecular complexes by multi-wavelength single-molecule fluorescence. *Biophys. J.* **91**, 1023–1031 (2006).
10. Sofia, S. J., Premnath, V. & Merrill, E. W. Poly(ethylene oxide) grafted to silicon surfaces: Grafting density and protein adsorption. *Macromolecules* **31**, 5059–5070 (1998).
11. Buschmann, V., Weston, K. D. & Sauer, M. Spectroscopic study and evaluation of red-absorbing fluorescent dyes. *Bioconjug. Chem.* **14**, 195–204 (2003).
12. Cordes, T., Maiser, A., Steinhauer, C., Schermelleh, L. & Tinnefeld, P. Mechanisms and advancement of antifading agents for fluorescence microscopy and single-molecule spectroscopy. *Phys. Chem. Chem. Phys.* **13**, 6699–6709 (2011).
13. Kim, Y. *et al.* Efficient site-specific labeling of proteins via cysteines.

Bioconjug. Chem. **19**, 786–791 (2008).

14. Yin, J. *et al.* Genetically encoded short peptide tag for versatile protein labeling by Sfp phosphopantetheinyl transferase. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 15815–15820 (2005).
15. Teh, Y. W., Jordan, M. I., Beal, M. J. & Blei, D. M. Hierarchical Dirichlet processes. *J. Am. Stat. Assoc.* **101**, 1566–1581 (2006).
16. Van Gael, J., Saatci, Y., Teh, Y. W. & Ghahramani, Z. Beam sampling for the infinite hidden Markov model. *Proc. 25th Int. Conf. Mach. Learn.* 1088–1095 (2008). doi:10.1145/1390156.1390293
17. Kuchta, R. D., Mizrahi, V., Benkovic, P. A., Johnson, K. A. & Benkovic, S. J. Kinetic Mechanism of DNA Polymerase I (Klenow). *Biochemistry* **26**, 8410–8417 (1987).
18. Christian, T. D., Romano, L. J. & Rueda, D. Single-molecule measurements of synthesis by DNA polymerase with base-pair resolution. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 21109–21114 (2009).
19. Markiewicz, R. P., Vrtis, K. B., Rueda, D. & Romano, L. J. Single-molecule microscopy reveals new insights into nucleotide selection by DNA polymerase i. *Nucleic Acids Res.* **40**, 7975–7984 (2012).
20. Wilkinson, A. *et al.* Analysis of ligation and DNA binding by Escherichia coli DNA ligase (LigA). *Biochim. Biophys. Acta - Proteins Proteomics* **1749**, 113–122 (2005).
21. Wang, H., Tang, Z., Wang, Y., Ma, G. & Tao, N. Probing Single Molecule Binding and Free Energy Profile with Plasmonic Imaging of Nanoparticles. *J. Am. Chem. Soc.* **141**, 16071–16078 (2019).
22. Friedman, L. J. & Gelles, J. Multi-wavelength single-molecule fluorescence analysis of transcription mechanisms. *Methods* **86**, 27–36 (2015).
23. Zhao, G., Gleave, E. S. & Lamers, M. H. Single-molecule studies contrast ordered DNA replication with stochastic translesion synthesis. *Elife* 1–19 (2017).
24. Ha, T. *et al.* Ligand-induced conformational changes observed in single RNA molecules. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 9077–9082 (1999).
25. Bieling, P., Telley, I. A., Hentrich, C., Piehler, J. & Surrey, T. *Fluorescence microscopy assays on chemically functionalized surfaces for quantitative imaging of microtubule, motor, and +TIP dynamics*. *Methods in Cell Biology* **95**, (Elsevier, 2010).
26. Chandradoss, S. D. *et al.* Surface passivation for single-molecule protein studies. *J. Vis. Exp.* 1–8 (2014).
27. Peterson, E. M., Manhart, M. W. & Harris, J. M. Single-Molecule

Fluorescence Imaging of Interfacial DNA Hybridization Kinetics at Selective Capture Surfaces. *Anal. Chem.* **88**, 1345–1354 (2016).

28. Oleksiievets, N. *et al.* Wide-Field Fluorescence Lifetime Imaging of Single Molecules. *J. Phys. Chem. A* **124**, 3494–3500 (2020).
29. George, N., Pick, H., Vogel, H., Johnsson, N. & Johnsson, K. Specific labeling of cell surface proteins with chemically diverse compounds. *J. Am. Chem. Soc.* **126**, 8896–8897 (2004).
30. La Clair, J. J., Foley, T. L. & Schegg, T. Manipulation of Carrier Proteins in Antibiotic Biosynthesis. *Chem. Biol.* **11**, ARTMED1118 (2004).
31. Yin, J., Liu, F., Li, X. & Walsh, C. T. Labeling proteins with small molecules by site-specific posttranslational modification. *J. Am. Chem. Soc.* **126**, 7754–7755 (2004).