



**HAL**  
open science

# A Toolbox for Site-Specific Labeling of RecQ Helicase With a Single Fluorophore Used in the Single-Molecule Assay

Fang-Yuan Teng, Zong-Zhe Jiang, Ling-Yun Huang, Man Guo, Feng Chen, Xi-Miao Hou, Xu-Guang Xi, Yong Xu

► **To cite this version:**

Fang-Yuan Teng, Zong-Zhe Jiang, Ling-Yun Huang, Man Guo, Feng Chen, et al.. A Toolbox for Site-Specific Labeling of RecQ Helicase With a Single Fluorophore Used in the Single-Molecule Assay. *Frontiers in Molecular Biosciences*, 2020, 7, 10.3389/fmolb.2020.586450 . hal-03080997

**HAL Id: hal-03080997**

**<https://hal.science/hal-03080997>**

Submitted on 18 Dec 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



# A Toolbox for Site-Specific Labeling of RecQ Helicase With a Single Fluorophore Used in the Single-Molecule Assay

Fang-Yuan Teng<sup>1,2,3,4†</sup>, Zong-Zhe Jiang<sup>1,3†</sup>, Ling-Yun Huang<sup>2</sup>, Man Guo<sup>3</sup>, Feng Chen<sup>1</sup>, Xi-Miao Hou<sup>2</sup>, Xu-Guang Xi<sup>2,5\*</sup> and Yong Xu<sup>1,3\*</sup>

<sup>1</sup> Experimental Medicine Center, The Affiliated Hospital of Southwest Medical University, Luzhou, China, <sup>2</sup> State Key Laboratory of Crop Stress Biology for Arid Areas and College of Life Sciences, Northwest A&F University, Yangling, China, <sup>3</sup> Department of Endocrinology and Metabolism, and Cardiovascular and Metabolic Diseases Key Laboratory of Luzhou, and Sichuan Clinical Research Center for Nephropathy, The Affiliated Hospital of Southwest Medical University, Luzhou, China, <sup>4</sup> Academician (Expert) Workstation of Sichuan Province, The Affiliated Hospital of Southwest Medical University, Luzhou, China, <sup>5</sup> LBPA, Ecole normale supérieure Paris-Saclay, Centre national de la recherche scientifique (CNRS), Université Paris Saclay, Cachan, France

## OPEN ACCESS

### Edited by:

Liang Hong,  
Shanghai Jiao Tong University, China

### Reviewed by:

Qin Xu,  
Shanghai Jiao Tong University, China  
Jianting Zheng,  
Shanghai Jiao Tong University, China

### \*Correspondence:

Xu-Guang Xi  
xxi01@ens-cachan.fr  
Yong Xu  
xywyll@aliyun.com

† These authors have contributed  
equally to this work

### Specialty section:

This article was submitted to  
Biophysics,  
a section of the journal  
Frontiers in Molecular Biosciences

Received: 23 July 2020

Accepted: 25 August 2020

Published: 25 September 2020

### Citation:

Teng F-Y, Jiang Z-Z, Huang L-Y,  
Guo M, Chen F, Hou X-M, Xi X-G and  
Xu Y (2020) A Toolbox for  
Site-Specific Labeling of RecQ  
Helicase With a Single Fluorophore  
Used in the Single-Molecule Assay.  
Front. Mol. Biosci. 7:586450.  
doi: 10.3389/fmolb.2020.586450

Fluorescently labeled proteins can improve the detection sensitivity and have been widely used in a variety of biological measurements. In single-molecule assays, site-specific labeling of proteins enables the visualization of molecular interactions, conformational changes in proteins, and enzymatic activity. In this study, based on a flexible linker in the *Escherichia coli* RecQ helicase, we established a scheme involving a combination of fluorophore labeling and sortase A ligation to allow site-specific labeling of the HRDC domain of RecQ with a single Cy5 fluorophore, without inletting extra fluorescent domain or peptide fragment. Using single-molecule fluorescence resonance energy transfer, we visualized that Cy5-labeled HRDC could directly interact with RecA domains and could bind to both the 3' and 5' ends of the overhang DNA dynamically *in vitro* for the first time. The present work not only reveals the functional mechanism of the HRDC domain, but also provides a feasible method for site-specific labeling of a domain with a single fluorophore used in single-molecule assays.

**Keywords:** fluorescence, molecular interaction, molecular dynamic, DNA repair, single molecule, helicase, protein labeling

## INTRODUCTION

Fluorescently labeled proteins can help dissect the detailed molecular mechanisms and have become crucial experimental tools in various research fields, especially in flow cytometry, fluorescence microscopy, and enzymatic activity measurements (Li et al., 2015). There are a variety of fluorescent labeling methods and strategies that depend on particular applications, such as genetically encoded tags (Schlichthaerle et al., 2019), quantum dots (Abu-Thabit and Ratemi, 2020), and small organic fluorophores (Rosen and Francis, 2017). It can be performed in solution and in real time, with high time resolution and high sensitivity even at the single-molecule level (Toseland, 2013).

In single-molecule fluorescence resonance energy transfer (smFRET) assay, fluorescently labeled proteins can realize the visualization of molecular interactions, conformational changes in proteins, and enzymatic activity and the ability to detect individual proteins moving in real time

(Aggarwal and Ha, 2016). In contrast to other applications, fluorescence labeling used in smFRET requires high precision because of its nanolevel sensitivity. Therefore, small organic fluorophores, such as NHS ester-activated cyanine, which can selectively couple with the N-terminus and lysine residues (Rosen and Francis, 2017), and maleimide-activated cyanine, which can selectively couple with cysteine residues (Friedman Ohana et al., 2019), have been mainly adopted to label the protein in smFRET because they are small and stable and do not perturb the real-time detection system. However, in the smFRET assay, it is important that the structure and activity of the protein will not be affected in fluorescence labeling process, and the fluorescence labeling site is specific, single, and homogeneous. Therefore, the specific strategy to label a small fluorophore to the indicated site plays a decisive role and becomes a key limiting factor in many smFRET detections.

RecQ family helicases include *Escherichia coli* RecQ, and *Homo sapiens* RecQ1, BLM, WRN, RecQ4, and RecQ5 $\beta$ , which widely participate in multiple DNA metabolic processes and play key roles in maintaining genomic stability (Bell and Kowalczykowski, 2016; de Renty and Ellis, 2017). As a functional and structural prototype of RecQ family helicases, *E. coli* RecQ can help suppress illegitimate recombination, repair stalled replication forks at DNA damage sites, and unwind G-quadruplex DNA (Croteau et al., 2014; Mendoza et al., 2016). Two conserved domains have been identified among all RecQ family helicases: the helicase domain, which contains two RecA domains and is involved in ssDNA binding and NTP hydrolysis, and the RecQ-C-terminal (RQC) domain, which is primarily responsible for substrate recognition and DNA unwinding (Figure 1A; Bernstein et al., 2003; Swan et al., 2014; Newman et al., 2015).

Additionally, some RecQ family helicases, including *E. coli* RecQ, and *H. sapiens* BLM and WRN, contain an auxiliary domain: the helicase and RNaseD C-terminal (HRDC) domain, which is connected to the RQC domain by a flexible linker [more than 20 amino acids (aa)] (Morozov et al., 1997). The HRDC domain of BLM has been reported to possess markedly low ssDNA-binding affinity ( $K_D \sim 100 \mu\text{M}$ ) (Kim and Choi, 2010) and interacts directly with the two RecA domains, which may influence the mechanochemical coupling of the ATPase cycle (Swan et al., 2014; Newman et al., 2015). However, the HRDC of human WRN has been shown to lack DNA binding ability *in vitro*, and it is suggested that it may mediate protein-protein interactions (Kitano et al., 2007). Meanwhile, the isolated HRDC of *E. coli* RecQ could bind only to ssDNA by electrophoretic mobility shift assay analysis (Bernstein and Keck, 2005). Enzyme kinetics studies found that the HRDC domain could slow down ssDNA translocation and dsDNA unwinding processes (Kocsis et al., 2014; Harami et al., 2015). Therefore, HRDC from different proteins may have different functions, and direct visualization of the trajectory of HRDC during RecQ transaction with different DNA at the single-molecule level needs to be detected.

In this study, based on the flexible linker ( $\sim 22$  aa) between the RQC and HRDC domain of *E. coli* RecQ helicase, we established a scheme involving a combination of fluorophore labeling and ligation by sortase A, which could fuse an LPXTG recognition

motif to an N-terminal GGG motif, thus regenerating a native amide bond and a recombinant protein (Levary et al., 2011). Therefore, we replaced only six non-functional amino acids on the flexible linker without insetting extra fluorescent domain or peptide fragment, succeeded in site-specific labeling of the HRDC domain of *E. coli* RecQ with a single Cy5 fluorophore, and found that the enzymatic activity of RecQ after fluorophore-labeling had little impact on the smFRET assay. Meanwhile, we observed that the HRDC domain could directly interact with RecA domains and could bind to both the 3' and 5' ends of the overhang DNA repeatedly. The present work not only directly reveals the functional mechanism of the HRDC domain during *E. coli* RecQ transaction with different DNA during DNA repair or DNA recombination, but also provides a feasible method for site-specific labeling of a domain with a single fluorophore used in single-molecule assays.

## MATERIALS AND METHODS

### Plasmid Construction, Protein Expression, and Purification

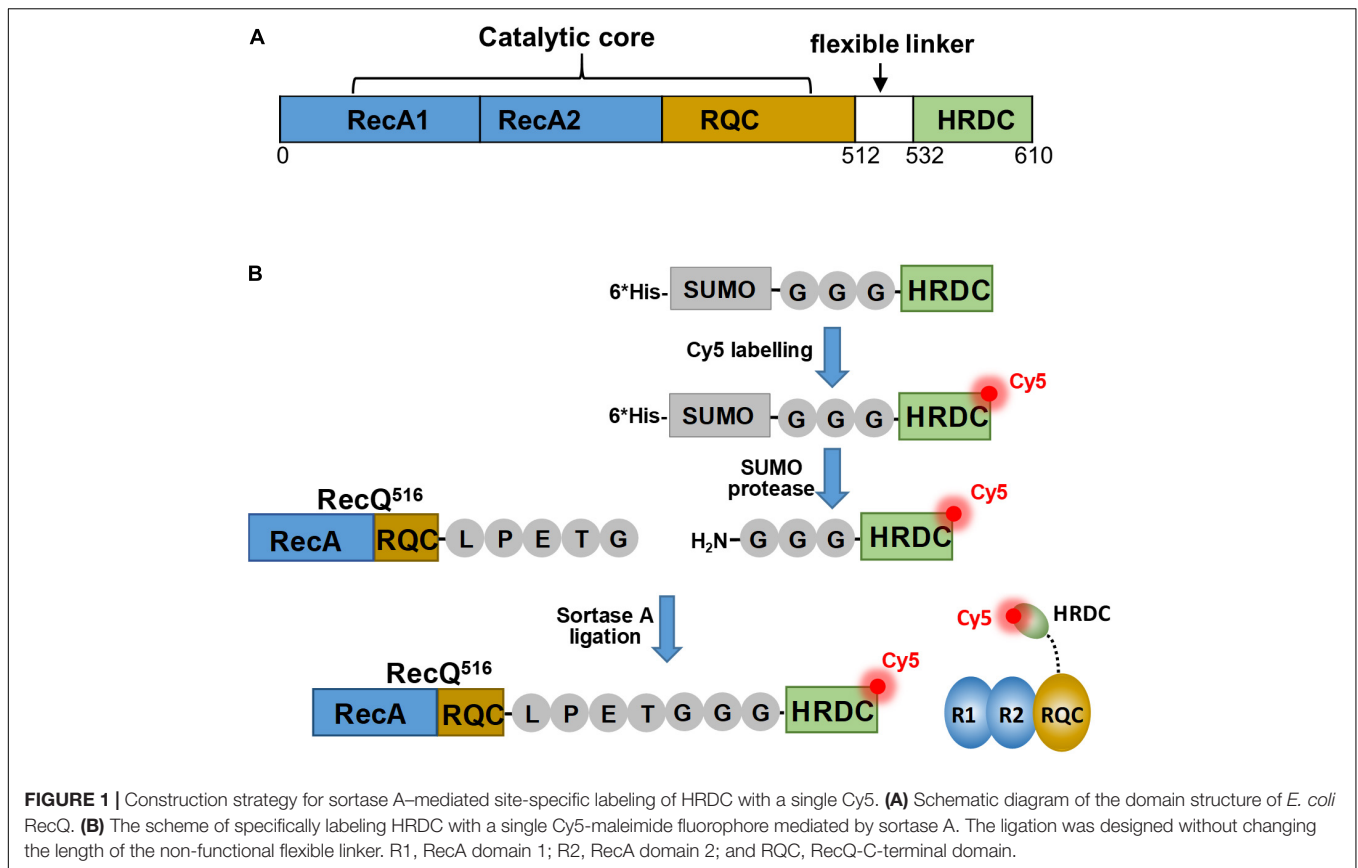
All DNA primers required making the protein constructs were purchased from Sangon Biotech (Shanghai, China), and the sequences were listed in **Supplementary Table S1**. All the indicated protein constructs were severally amplified from *E. coli* genome. The sumo-GGG-HRDC construct was obtained by overlap polymerase chain reaction (PCR) using the primer Sumo-F/Sumo-R, and HRDC-F/HRDC-R. RecQ<sup>516</sup>-LPETG construct was obtained by PCR using the primer RecQ<sup>516</sup>-LPETG-F and RecQ<sup>516</sup>-LPETG-R. After digestion by *NdeI/XhoI*, the indicated protein constructs were severally constructed into pET15b vector, and expressed in BL21 (DE3) induced by 0.3 mM IPTG at 18°C for 16 h. Then, the indicated recombinant protein was purified by Ni affinity chromatography (Shi et al., 2017). Briefly, after being harvested by centrifugation, each cell pellet was resuspended by ice-cold lysis buffer (10 mM imidazole and 500 mM NaCl in 20 mM Tris-HCl, pH 8.0), crushed with a French press and ultrasonicated for three turns. Subsequently, the supernatants were separated by centrifugation at 12,000 g for 30 min at 4°C and then loaded into Ni affinity chromatography column (GE Healthcare, Chicago, IL, United States). After being washed by 20 column volumes of lysis buffer containing 30 mM imidazole, the purified protein was eluted from Ni affinity resin by elution buffer (300 mM imidazole and 500 mM NaCl in 20 mM Tris-HCl, pH 8.0).

### Sortase A-Mediated Ligation Reaction

The sortase A-mediated ligation reaction was in ligation buffer containing 50 mM Tris-HCl, pH 7.0, 150 mM NaCl, and 20 mM CaCl<sub>2</sub>. The ligation was conducted with 20  $\mu\text{M}$  RecQ<sup>516</sup>-LPETG, 100  $\mu\text{M}$  Cy5-labeled GGG-HRDC, and 50  $\mu\text{M}$  sortase A at 34°C for 1 h.

### RecQ Reaction Buffer

RecQ reaction buffer contains 2 mM MgCl<sub>2</sub>, 0.1 mg/mL bovine serum albumin, and 50 mM KCl in 20 mM Tris-HCl, pH 7.5.



For single-molecule measurements, 0.8% D-glucose, 1 mg/mL glucose oxidase (266,600 units/g, Sigma-Aldrich, St. Louis, MO, United States), 0.4 mg/mL catalase (2,000–5,000 units/mg, Sigma-Aldrich), and 1 mM Trolox (Sigma-Aldrich) were added to the reaction buffer (Roy et al., 2008).

## Single-Molecule Fluorescence Data Acquisition

All oligonucleotides required to make DNA substrates were purchased from Sangon Biotech. DNA constructs used in single-molecule measurements were carried out as described previously (Wang et al., 2019), and the sequences were listed in **Supplementary Table S2**. SmFRET study was carried out with a home-built objective-type total-internal-reflection microscopy as described previously (Roy et al., 2008; Wang et al., 2019). Cy3 was excited by a 532-nm Sapphire laser (Coherent, Santa Clara, CA, United States). An oil immersion objective (100×, N.A.1.49) was used to generate an evanescent field of illumination. Fluorescence signals from Cy3 and Cy5 were split by a dichroic mirror and finally collected by an electron-multiplying charge-coupled device camera (iXON; Andor Technology, South Windsor, CT, United States). Fluorescence imaging processes were controlled and recorded by MetaMorph (Molecular Device, Sunnyvale, CA, United States). The coverslips (Fisher Scientific, Pittsburgh, PA, United States) and slides were cleaned thoroughly by a mixture of sulfuric acid and hydrogen peroxide, acetone, and

sodium ethoxide, and then the surfaces of coverslip were coated with a mixture of 99% mPEG (m-PEG-5000, Laysan Bio, Inc., Arab, AL, United States) and 1% of biotin-PEG (biotin-PEG-5000, Laysan Bio, Inc.). Streptavidin (10 μg/mL) was added to the microfluidic chamber made of the PEG-coated coverslip and incubated for 10 min. After washing, 100 pM DNA was immobilized for 10 min. Then free DNA was removed by washing with the reaction buffer. We used an exposure time of 100 ms for all single-molecule measurements at a constant temperature of 22°C. To obtain the fraction of DNA unwinding vs. time, a series of movies were recorded with 1-s duration at indicated times, and the Cy3 spots were counted to represent the number of remaining DNA molecules.

## FRET Data Analyses

The FRET efficiency was calculated using  $I_A/(I_D + I_A)$ , where  $I_D$  and  $I_A$  represent the intensity of donor and acceptor, respectively. The leakage from Cy3 Channel to Cy5 Channel is about 10%; therefore, we deducted the leakage when exporting the single-molecule fluorescence intensity by the software “smCamera” (Roy et al., 2008). Basic data analysis was carried out by scripts written in MATLAB, and all data fitting was generated by Origin 8.0. An automated step-finding method (from <http://bio.physics.illinois.edu/HaMMMy.asp>) was employed to characterize the association and dissociation of RecQ, and the  $E_{\text{FRET}}$  value and dwell time ( $t$ ) for each reaction were determined accordingly. The resulting histograms of FRET values



and dwell time from more than 150 molecules were fitted with multipeak Gaussian distribution or Gamma distribution or single-exponential decay, respectively.

## RESULTS

### Labeling the HRDC Domain With a Single Cy5 Fluorophore Mediated by Sortase A Ligation

Maleimide-activated fluorophores are monoreactive dyes that can selectively couple with cysteines in peptides or proteins to generate specifically labeled conjugates and have been widely used in enzymology experiments. As full-length RecQ contains 11 cysteine residues, it is difficult to label the HRDC domain of RecQ with a single fluorophore directly. To avoid non-specific fluorescent labeling, based on the flexible linker (~22 aa) between the RQC and HRDC domains, we established a scheme (**Figure 1B**) to specifically label the HRDC domain with a single Cy5-maleimide fluorophore mediated by sortase A, which could fuse an LPXTG recognition motif to an N-terminal GGG-containing motif, thus regenerating a native amide bond and a recombinant protein.

First, owing to the absence of cysteine in the HRDC domain, E610C was designed to label the Cy5-maleimide fluorophore. To obtain sortase A recognition sequence H<sub>2</sub>N-GGG, we conducted overlap PCR to fuse HRDC (amino acids 524–610) with a SUMO tag (12 kDa) and amino acid sequence “GGG” at the N-terminal (**Figure 1B**). The recombinant sequence was cloned into pET15b and expressed in BL21 (DE3). Then, the recombinant protein sumo-GGG-HRDC (~21 kDa) was purified by Ni affinity chromatography (**Figure 2A**). After that, sumo-GGG-HRDC was labeled with a 15-fold molar excess of Cy5-maleimide fluorophore at 20°C for 1 h, and the free Cy5 dye was removed by Ni affinity chromatography. Then, the Cy5-labeled sample was digested with SUMO protease at 4°C overnight

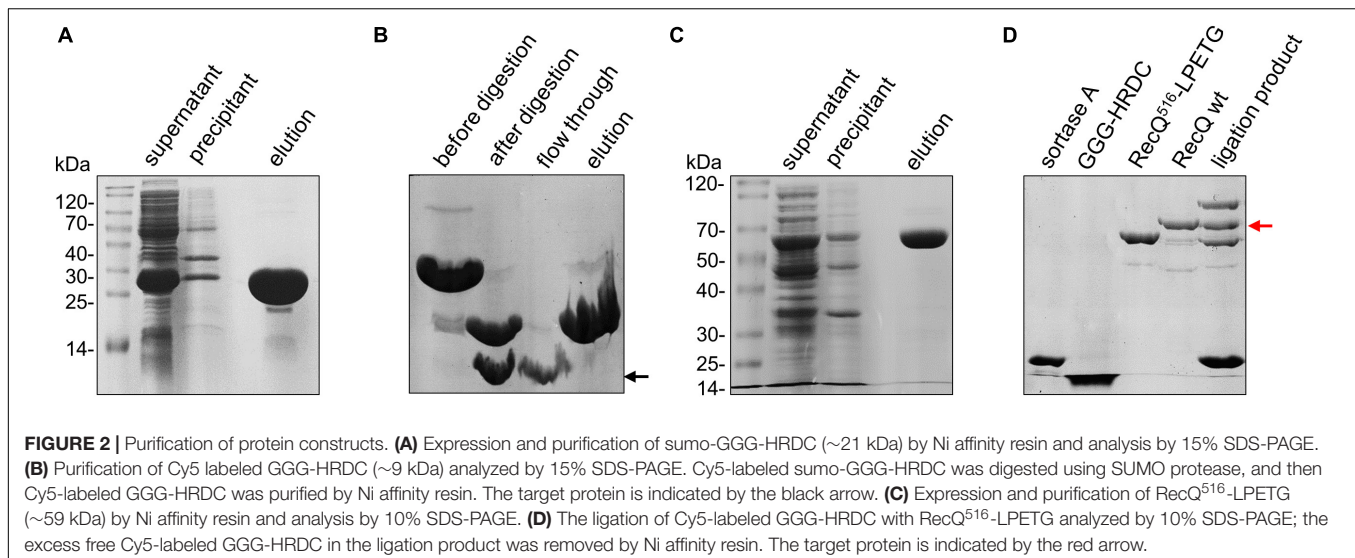
(**Figure 2B**). Finally, the Cy5-labeled GGG-HRDC (~9 kDa) was purified again by Ni affinity chromatography (indicated by black arrow in **Figure 2B**). We then used the preset-program “Protein and Labels” of Thermo Scientific Nanodrop 2000c to measure the concentration of protein and Cy5. As the Cy5 labeling was single and site-specific, and the free Cy5 dye was removed, the labeling efficiency was calculated by the ratio of the Cy5 concentration to the protein concentration, and it was more than 95%.

In the meantime, we conducted overlap PCR to fuse sortase A recognition sequence “LPETG” at the C-terminal of RecQ core (amino acids 1–516, designated as RecQ<sup>516</sup>-LPETG, ~59 kDa). RecQ<sup>516</sup>-LPETG was cloned into pET15b and expressed in BL21 (DE3). Similarly, RecQ<sup>516</sup>-LPETG was purified by Ni affinity chromatography (**Figure 2C**). Additionally, the expression and purification of sortase A (~19 kDa) were carried out as described previously (Mao et al., 2004).

The sortase A-mediated ligation reaction was conducted with 20 μM RecQ<sup>516</sup>-LPETG, 100 μM Cy5-labeled GGG-HRDC, and 50 μM sortase A in ligation buffer (50 mM Tris-HCl, pH 7.0, 150 mM NaCl, and 20 mM CaCl<sub>2</sub>) at 34°C for 1 h. At last, Ni affinity chromatography was used again to remove the excess free Cy5-labeled GGG-HRDC, as free Cy5-labeled GGG-HRDC will disturb the fluorescence signal; the ligation efficiency was ~30%, as determined by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) (**Figure 2D**). Therefore, we succeeded in labeling the HRDC domain of *E. coli* RecQ with a single Cy5 fluorophore, and the Cy5-labeled recombinational RecQ was designated as RecQ-Cy5.

### Cy5-Labeled RecQ Exhibits the Same Unwinding Activities as Those of Wild-Type RecQ

First, we designed the 10-nt overhang DNA substrate to evaluate the activity of RecQ-Cy5. The DNA substrate contained a Cy3 fluorophore labeled at the 3' end of the 10-nt overhang, and a biotin labeled at the 3' end of the complementary strand to fix the



DNA substrate on the surface of the reaction chamber, such that the Cy3 signal would disappear promptly if the 18-bp dsDNA was fully unwound (Figure 3A).

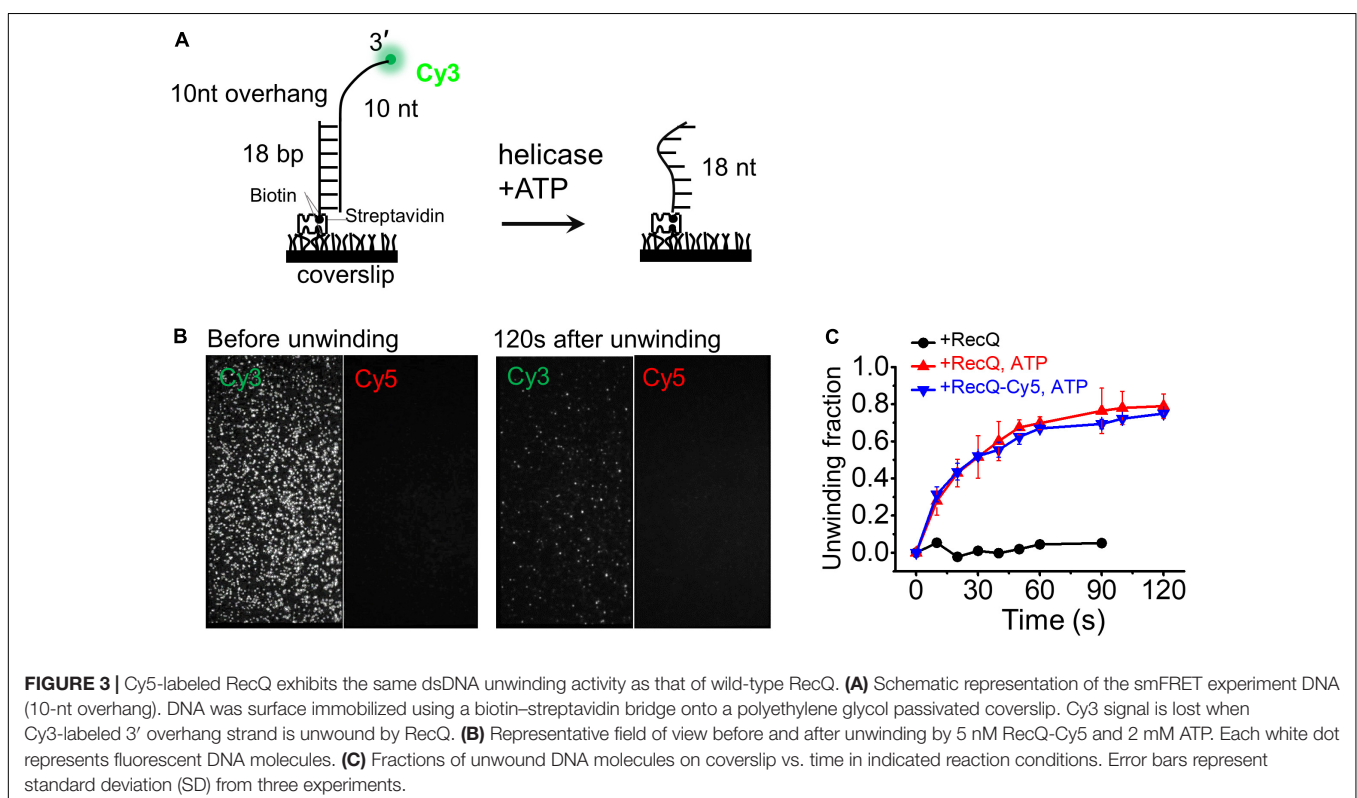
Then, we measured the unwinding activity and unwinding fractions of the 18-bp 10-nt DNA by 5 nM RecQ-Cy5 or wild-type RecQ using 2 mM ATP. To obtain the fraction of DNA unwinding vs. time, the fluorescence is intermittent excited, and a series of movies were recorded with 1-s duration at indicated times, and the Cy3 spots were counted to represent the number of remaining DNA molecules, as previously reported (Wang et al., 2018). We found that RecQ-Cy5 could unwind dsDNA (Figure 3B), and the unwinding fractions of 18-bp 10-nt DNA by RecQ-Cy5 and wild-type RecQ had little difference (Figure 3C). Therefore, we concluded that Cy5 labeling or sortase A ligation did not affect the RecQ-Cy5 activity, and RecQ-Cy5 occupied the same unwinding activities as wild-type RecQ.

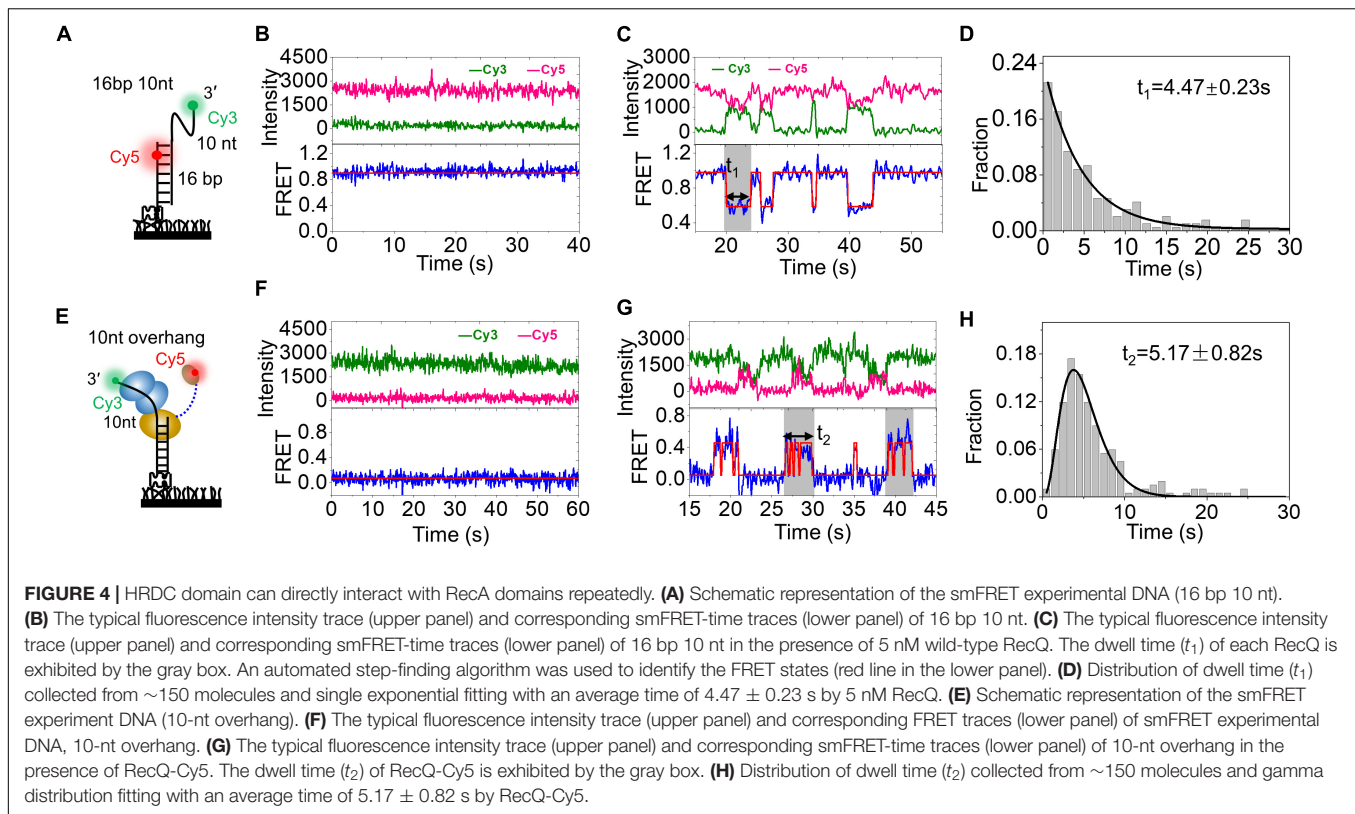
### The HRDC Domain Can Directly Interact With RecA Domains Repeatedly

Before detecting the dynamic trajectory of Cy5-labeled HRDC, we first detected the binding process of wild-type RecQ by smFRET. The DNA substrate 16 bp with 3' 10-nt overhang was prepared (Figure 4A). The donor (Cy3) was labeled at the 3' end of the 10-nt overhang, and the acceptor (Cy5) was labeled at the 4th nucleotide from the 5' end in the 16-nt stem strand so that a relatively high FRET efficiency ( $E_{\text{FRET}}$ ) of  $\sim 0.92$  was detected because of the high flexibility of ssDNA (Figure 4B and Supplementary Figure S1A). Upon addition of wild-type RecQ, smFRET-time traces exhibited periodic fluctuations between an

$E_{\text{FRET}}$  of  $\sim 0.92$  and  $\sim 0.54$  (Figure 4C and Supplementary Figure S1A), as RecQ-binding stretched ssDNA. Then, we analyzed the dwell time of the periodic fluctuations, which has been indicated by the gray box and represents RecQ bound to the DNA. The dwell time  $t_1$  collected from more than 150 molecules and followed single-exponential decay with an average time of  $4.47 \pm 0.23$  s (Figure 4D).

We then detected the dynamic trajectory of RecQ-Cy5 binding to the 10-nt overhang DNA substrate, in which the donor (Cy3) was labeled at the 3' end of the overhang (Figure 4E). The FRET efficiency of the substrate was  $\sim 0$  because of the lack of the acceptor (Cy5) (Figure 4F and Supplementary Figure S2A). After allowing RecQ-Cy5 (5 nM) binding to substrate DNA for 2 min, we washed off the free protein in the reaction chamber using reaction buffer and recorded the binding process. The binding process exhibited rapid periodic fluctuations between  $E_{\text{FRET}}$  of  $\sim 0$  and  $\sim 0.5$  within a certain period, as indicated by the gray box in Figure 4G. The total duration time of each rapid periodic fluctuation ( $t_2$  in Figure 4G) was counted, and it followed the gamma distribution (Zhou et al., 2014; Wu et al., 2015) with an average time of  $5.17 \pm 0.82$  s (Figure 4H), which is comparable to the dwell time ( $t_1 = 4.47 \pm 0.23$  s) when wild-type RecQ was binding to the 3' overhang DNA, as shown in Figure 4C; this result indicated that the rapid periodic fluctuations in  $t_2$  were caused by HRDC rapid periodic association to and dissociation from RecA domains after the binding of RecQ-Cy5 to substrate DNA, rather than RecQ repeatedly binding to DNA.





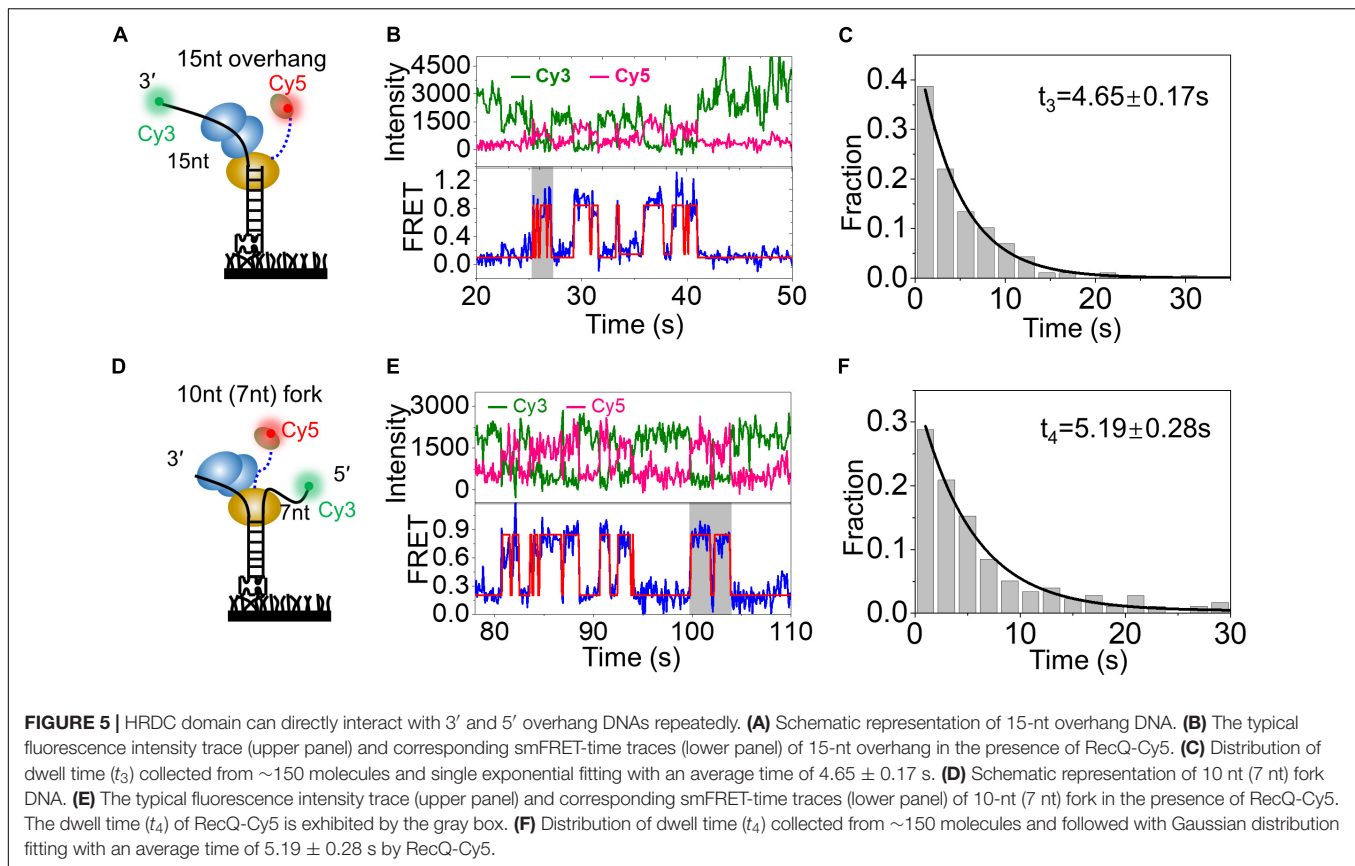
We constructed a rough structure of *E. coli* RecQ in complex with a partial duplex according to the existing structural data (Supplementary Figure S2B). The simulated structure presented that RecA domain of RecQ would bind with  $\sim 9$ -nt ssDNA, and HRDC domain may directly interact with RecA domain. Meanwhile, the C-terminal of HRDC domain from RecQ was just  $\sim 6.2$  nm away from the 10th base of the 3' overhang DNA (Supplementary Figure S2B), which could exactly give rise to a FRET at  $\sim 0.45$ . Therefore, under our experimental conditions, the rapid periodic fluctuations of FRET between  $\sim 0$  and  $\sim 0.49$ , as indicated by the gray box in Figure 4G, could be caused by the interaction of the Cy5-labeled HRDC domain with RecA domains repeatedly after RecQ binding to DNA. Therefore, we concluded that the site-specific fluorescent labeling method established by us could visualize the interaction between molecules, and HRDC would repeatedly interact with RecA domains when RecQ binds to 3' overhang DNA.

### The HRDC Domain Can Directly Interact With the 3' and 5' Overhang DNA Repeatedly

Then, a substrate DNA with 15-nt 3' overhang was designed, and the donor (Cy3) was also labeled at the 3' end of the overhang (Figure 5A). After RecQ-Cy5 (5 nM) binding to substrate DNA for 2 min and when the free protein was washed off by reaction buffer, FRET rapidly increased from  $\sim 0$  to  $\sim 0.84$  repeatedly, as indicated in the gray box in Figure 5B

(Figure 5B and Supplementary Figure S3A); this was much higher than that of 10-nt 3' overhang DNA, corresponding to RecQ binding to substrate DNA, and Cy5-labeled HRDC domain much closer to the 3' end of the 15-nt overhang than the 10-nt overhang DNA. The average duration of rapid periodic fluctuation ( $t_3$ ) was counted from more than 150 molecules and followed by single exponential decay fitting with an average time of  $4.65 \pm 0.17$  s (Figure 5C), which was also comparable to the dwell time when wild-type RecQ was binding to 3' overhang DNA (Figure 4D). These results indicate that when there is additional ssDNA (more than 10 nt) in the 3' end of the overhang, the Cy5-labeled HRDC domain can rapidly and repeatedly bind to the 3' end of the overhang directly.

In the meantime, a fork DNA with 10-nt 3' overhang and 7-nt 5' overhang was designed (Figure 5D), and the donor (Cy3) was labeled at the 5' end of the overhang. Under the same conditions, FRET also rapidly increased from  $\sim 0$  to  $\sim 0.92$  repeatedly (Figure 5E and Supplementary Figure S3B), thereby indicating that Cy5 was almost completely close to Cy3. The average duration of rapid periodic fluctuation ( $t_4$ ) was also determined.  $t_4$  also followed single exponential decay fitting with an average duration of  $5.19 \pm 0.28$  s (Figure 5F). The aforementioned results indicated that the Cy5-labeled HRDC domain could also bind to the 5' end of overhang repeatedly after RecQ binding to the substrate DNA. Based on these data, we concluded that HRDC could directly interact with 3' and 5' overhang DNA repeatedly after RecQ binding to the substrate DNA.



## DISCUSSION

Fluorescent dyes, such as cyanine and Alexa Fluor, have been widely used in *in vitro* enzymology experiments because they are easy to label and have little effect on enzymatic activity (Xi and Deprez, 2010). NHS ester-activated and maleimide-activated dyes have been widely adopted and can be specifically coupled with the amine group of lysine residues or the N-terminus and the sulfhydryl of cysteine residues, respectively. However, as lysine residues are more abundant in proteins (average abundance, 5.9%) and are frequently involved in binding interactions (Rosen and Francis, 2017), cysteine residues (average abundance, 1.9%) always stand out as uniquely reactive sites for labeling fluorescent dyes to detect conformational changes or trajectory of proteins in smFRET experiments (Christian et al., 2009; Hou et al., 2015; Bell and Kowalczykowski, 2016). Usually, to label one fluorophore to a definite cysteine site, other cysteine residues must be mutated. Meanwhile, it should ensure that the structure and activity of the protein will not be affected when mutating cysteine residues. Therefore, fluorophore labeling of proteins is a key limiting factor in many smFRET experiments.

*E. coli* RecQ contains 11 cysteine residues, and it is difficult to label the HRDC domain with fluorophore directly as it is almost impossible to mutate all cysteines without affecting the structure and activity of RecQ (Ren et al., 2008). Therefore, we established a scheme for specifically labeling HRDC with a single Cy5-maleimide fluorophore, which is mediated by sortase A and

is dependent on the flexible linker ( $\sim 22$  aa) between the RQC and HRDC domains (Figure 1).

Sortase A is an efficient and versatile tool for protein modification, which can fuse an LPXTG recognition motif to an N-terminal GGG-containing motif, thereby regenerating a native amide bond (Li et al., 2020). Therefore, our scheme is mainly appropriate for proteins that contain multiple or immutable cysteines on the non-labeled domain, and there is a non-functional flexible linker between the labeled domain and non-labeled domain for recognition and ligation by sortase A. Meanwhile, it is worth noting that our scheme replaces only six non-functional amino acids on the flexible linker and will not inlet extra fluorescent domain or peptide fragment, which is highly likely to affect the protein activity or structure, thereby leading to perturbations in the real-time detection system because of the high sensitivity of smFRET.

The ligation efficiency of sortase A is affected by many factors, such as the steric hindrance of the proligation motifs, ligation buffer, and temperature (Broguiere et al., 2018). Usually, a low-salt solution (150 mM  $\text{Na}^+$ ), appropriate  $\text{Ca}^{2+}$  concentration (5–60 mM), and proper temperature ( $\sim 37^\circ\text{C}$ ) can increase the ligation efficiency of sortase A (Dasgupta et al., 2011; Levary et al., 2011; Warden-Rothman et al., 2013; Broguiere et al., 2018). Meanwhile, it should be noted that low salt solution and high temperature may cause protein denaturation during ligation. Certainly, apart from sortase A-mediated ligation, other ligation



methods may also be feasible, such as protein trans-splicing by split inteins (Friedel et al., 2019; Yao et al., 2020).

We visualized that Cy5-labeled HRDC domain could repeatedly interact with RecA domains and bind to 3' and 5' overhang DNAs repeatedly after RecQ binding to DNA for the first time (Figures 4, 5). Repeated interaction of HRDC with RecA domains may inhibit ATP binding or ATP hydrolysis, thereby affecting ADP, and/or Pi release by RecA, and thus slowing down the ssDNA translocation and dsDNA unwinding processes, as previously reported (Chatterjee et al., 2014; Kocsis et al., 2014; Harami et al., 2015). The HRDC domain directly binds to the overhang DNA and may lead to multiple functions. While binding to the 3' overhang DNA, HRDC may directly slow down the ssDNA translocation or dsDNA unwinding, as previously reported (Kocsis et al., 2014; Harami et al., 2015). Meanwhile, by binding to 5' overhang DNA, HRDC may induce pausing and shuttling during unwinding dsDNA (Harami et al., 2017). Furthermore, we have also discovered that HRDC can facilitate strand-switch process and restrain RecQ patrolling on 5' ssDNA after HRDC establishes contact with the 5' overhang DNA (Teng et al., in preparation).

Owing to their low sequence identity and highly different surface properties, the HRDC domain from different RecQ family helicases may exhibit differentiated functions, especially the DNA binding and protein-protein interaction activities (Brosh, 2013), which may need to be detected at the single-molecule level in future studies. Our research not only directly reveals the functional mechanism of the HRDC domain during *E. coli* RecQ transaction with different DNA during DNA repair and DNA recombination, but also provides a feasible method for site-specific labeling of a domain with a single fluorophore and thus facilitates the visualization of molecular interactions, conformational changes in proteins, enzymatic activity, and individual protein moving in real time at the single-molecule level.

## REFERENCES

- Abu-Thabit, N., and Ratemi, E. (2020). Hybrid porous silicon biosensors using plasmonic and fluorescent nanomaterials: a mini review. *Front. Chem.* 8:454. doi: 10.3389/fchem.2020.00454
- Aggarwal, V., and Ha, T. (2016). Single-molecule fluorescence microscopy of native macromolecular complexes. *Curr. Opin. Struct. Biol.* 41, 225–232. doi: 10.1016/j.sbi.2016.09.006
- Bell, J. C., and Kowalczykowski, S. C. (2016). Mechanics and single-molecule interrogation of DNA recombination. *Annu. Rev. Biochem.* 85, 193–226. doi: 10.1146/annurev-biochem-060614-034352
- Bernstein, D. A., and Keck, J. L. (2005). Conferring substrate specificity to DNA helicases: role of the RecQ HRDC domain. *Structure* 13, 1173–1182. doi: 10.1016/j.str.2005.04.018
- Bernstein, D. A., Zittel, M. C., and Keck, J. L. (2003). High-resolution structure of the *E. coli* RecQ helicase catalytic core. *EMBO J.* 22, 4910–4921. doi: 10.1093/emboj/cdg500
- Broguere, N., Formica, F. A., Barreto, G., and Zenobi-Wong, M. (2018). Sortase A as a cross-linking enzyme in tissue engineering. *Acta Biomater.* 77, 182–190. doi: 10.1016/j.actbio.2018.07.020
- Brosh, R. M. Jr. (2013). DNA helicases involved in DNA repair and their roles in cancer. *Nat. Rev. Cancer* 13, 542–558. doi: 10.1038/nrc3560

## DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation, to any qualified researcher.

## AUTHOR CONTRIBUTIONS

X-GX and YX conceived and supervised the study. X-GX and F-YT designed experiments. F-YT, Z-ZJ, and L-YH performed experiments. F-YT and MG analyzed data. F-YT, Z-ZJ, and X-MH wrote the manuscript. X-GX, YX, and FC made manuscript revisions. All authors contributed to the article and approved the submitted version.

## FUNDING

This work was supported by the Research Startup Funding of the Affiliated Hospital of Southwest Medical University (Grant No. 18102), the Scientific Research Funding of Luzhou-Southwest Medical University (Grant Nos. 2019LZXNYDJ06 and 2018LZXNYD-PT01), the Scientific Research Funding of Southwest Medical University (Grant No. 2018-ZRZD-003), and the key projects of the Sichuan Science and Technology Department (Grant No. 2019YFS0537). The research was conducted within the context of the International Associated Laboratory “Helicase-mediated G-quadruplex DNA unwinding and Genome Stability.”

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmolb.2020.586450/full#supplementary-material>

- Chatterjee, S., Zigelbaum, J., Savitsky, P., Sturzenegger, A., Huttner, D., Janscak, P., et al. (2014). Mechanistic insight into the interaction of BLM helicase with intra-strand G-quadruplex structures. *Nat. Commun.* 5:5556. doi: 10.1038/ncomms6556
- Christian, T. D., Romano, L. J., and Rueda, D. (2009). Single-molecule measurements of synthesis by DNA polymerase with base-pair resolution. *Proc. Natl. Acad. Sci. U.S.A.* 106, 21109–21114. doi: 10.1073/pnas.0908640106
- Croteau, D. L., Popuri, V., Opresko, P. L., and Bohr, V. A. (2014). Human RecQ helicases in DNA repair, recombination, and replication. *Annu. Rev. Biochem.* 83, 519–552. doi: 10.1146/annurev-biochem-060713-035428
- Dasgupta, S., Samantaray, S., Sahal, D., and Roy, R. P. (2011). Isopeptide ligation catalyzed by quintessential sortase A: mechanistic cues from cyclic and branched oligomers of indolicidin. *J. Biol. Chem.* 286, 23996–24006. doi: 10.1074/jbc.M111.247650
- de Renty, C., and Ellis, N. A. (2017). Bloom's syndrome: why not premature aging? A comparison of the BLM and WRN helicases. *Ageing Res. Rev.* 33, 36–51. doi: 10.1016/j.arr.2016.05.010
- Friedel, K., Popp, M. A., Matern, J. C. J., Gazdag, E. M., Thiel, I. V., Volkmann, G., et al. (2019). A functional interplay between intein and extein sequences in protein splicing compensates for the essential block B histidine. *Chem. Sci.* 10, 239–251. doi: 10.1039/c8sc01074a

- Friedman Ohana, R., Hurst, R., Rosenblatt, M., Levin, S., Machleidt, T., Kirkland, T. A., et al. (2019). Utilizing a simple method for stoichiometric protein labeling to quantify antibody blockade. *Sci. Rep.* 9:7046. doi: 10.1038/s41598-019-43469-z
- Harami, G. M., Nagy, N. T., Martina, M., Neuman, K. C., and Kovacs, M. (2015). The HRDC domain of *E. coli* RecQ helicase controls single-stranded DNA translocation and double-stranded DNA unwinding rates without affecting mechanoenzymatic coupling. *Sci. Rep.* 5:11091. doi: 10.1038/srep11091
- Harami, G. M., Seol, Y., In, J., Ferenczi, V., Martina, M., Gyimesi, M., et al. (2017). Shuttling along DNA and directed processing of D-loops by RecQ helicase support quality control of homologous recombination. *Proc. Natl. Acad. Sci. U.S.A.* 114, E466–E475. doi: 10.1073/pnas.1615439114
- Hou, X. M., Wu, W. Q., Duan, X. L., Liu, N. N., Li, H. H., Fu, J., et al. (2015). Molecular mechanism of G-quadruplex unwinding helicase: sequential and repetitive unfolding of G-quadruplex by Pif1 helicase. *Biochem. J.* 466, 189–199. doi: 10.1042/Bj20140997
- Kim, Y. M., and Choi, B. S. (2010). Structure and function of the regulatory HRDC domain from human Bloom syndrome protein. *Nucleic Acids Res.* 38, 7764–7777. doi: 10.1093/nar/gkq586
- Kitano, K., Yoshihara, N., and Hakoshima, T. (2007). Crystal structure of the HRDC domain of human Werner syndrome protein, WRN. *J. Biol. Chem.* 282, 2717–2728. doi: 10.1074/jbc.M610142200
- Kocsis, Z. S., Sarlos, K., Harami, G. M., Martina, M., and Kovacs, M. (2014). A nucleotide-dependent and HRDC domain-dependent structural transition in DNA-bound RecQ helicase. *J. Biol. Chem.* 289, 5938–5949. doi: 10.1074/jbc.M113.530741
- Levary, D. A., Parthasarathy, R., Boder, E. T., and Ackerman, M. E. (2011). Protein-protein fusion catalyzed by sortase A. *PLoS One* 6:e18342. doi: 10.1371/journal.pone.0018342
- Li, J., Zhang, Y., Soubias, O., Khago, D., Chao, F. A., Li, Y., et al. (2020). Optimization of sortase A ligation for flexible engineering of complex protein systems. *J. Biol. Chem.* 295, 2664–2675. doi: 10.1074/jbc.RA119.012039
- Li, Z., Theile, C. S., Chen, G. Y., Bilate, A. M., Duarte, J. N., Avalos, A. M., et al. (2015). Fluorophore-conjugated holliday junctions for generating super-bright antibodies and antibody fragments. *Angew. Chem. Int. Ed. Engl.* 54, 11706–11710. doi: 10.1002/anie.201505277
- Mao, H., Hart, S. A., Schink, A., and Pollok, B. A. (2004). Sortase-mediated protein ligation: a new method for protein engineering. *J. Am. Chem. Soc.* 126, 2670–2671. doi: 10.1021/ja039915e
- Mendoza, O., Bourdoncle, A., Boule, J. B., Brosh, R. M. Jr., and Mergny, J. L. (2016). G-quadruplexes and helicases. *Nucleic Acids Res.* 44, 1989–2006. doi: 10.1093/nar/gkw079
- Morozov, V., Mushegian, A. R., Koonin, E. V., and Bork, P. (1997). A putative nucleic acid-binding domain in Bloom's and Werner's syndrome helicases. *Trends Biochem. Sci.* 22, 417–418. doi: 10.1016/S0968-0004(97)01128-6
- Newman, J. A., Savitsky, P., Allerston, C. K., Bizard, A. H., Ozer, O., Sarlos, K., et al. (2015). Crystal structure of the Bloom's syndrome helicase indicates a role for the HRDC domain in conformational changes. *Nucleic Acids Res.* 43, 5221–5235. doi: 10.1093/nar/gkv373
- Ren, H., Dou, S. X., Zhang, X. D., Wang, P. Y., Kanagaraj, R., Liu, J. L., et al. (2008). The zinc-binding motif of human RECQ5beta suppresses the intrinsic strand-annealing activity of its DExH helicase domain and is essential for the helicase activity of the enzyme. *Biochem. J.* 412, 425–433. doi: 10.1042/Bj20071150
- Rosen, C. B., and Francis, M. B. (2017). Targeting the N terminus for site-selective protein modification. *Nat. Chem. Biol.* 13, 697–705. doi: 10.1038/nchembio.2416
- Roy, R., Hohng, S., and Ha, T. (2008). A practical guide to single-molecule FRET. *Nat. Methods* 5, 507–516. doi: 10.1038/nmeth.1208
- Schlichthaerle, T., Strauss, M. T., Schueder, F., Auer, A., Nijmeijer, B., Kueblbeck, M., et al. (2019). Direct visualization of single nuclear pore complex proteins using genetically-encoded probes for DNA-PAINT. *Angew. Chem. Int. Ed.* 58, 13004–13008. doi: 10.1002/anie.201905685
- Shi, J., Liu, N. N., Yang, Y. T., and Xi, X. G. (2017). Purification and enzymatic characterization of Gallus gallus BLM helicase. *J. Biochem.* 162, 183–191. doi: 10.1093/jb/mvx013
- Swan, M. K., Legris, V., Tanner, A., Reaper, P. M., Vial, S., Bordas, R., et al. (2014). Structure of human Bloom's syndrome helicase in complex with ADP and duplex DNA. *Acta Crystallogr. Section D Biol. Crystallogr.* 70, 1465–1475. doi: 10.1107/S139900471400501x
- Toseland, C. P. (2013). Fluorescent labeling and modification of proteins. *J. Chem. Biol.* 6, 85–95. doi: 10.1007/s12154-013-0094-5
- Wang, L., Wang, Q. M., Wang, Y. R., Xi, X. G., and Hou, X. M. (2018). DNA-unwinding activity of *Saccharomyces cerevisiae* Pif1 is modulated by thermal stability, folding conformation, and loop lengths of G-quadruplex DNA. *J. Biol. Chem.* 293, 18504–18513. doi: 10.1074/jbc.RA118.005071
- Wang, Q. M., Yang, Y. T., Wang, Y. R., Gao, B., Xi, X. G., and Hou, X. M. (2019). Human replication protein A induces dynamic changes in single-stranded DNA and RNA structures. *J. Biol. Chem.* 294, 13915–13927. doi: 10.1074/jbc.RA119.009737
- Warden-Rothman, R., Caturegli, I., Popik, V., and Tsourkas, A. (2013). Sortase-tag expressed protein ligation: combining protein purification and site-specific bioconjugation into a single step. *Anal. Chem.* 85, 11090–11097. doi: 10.1021/ac402871k
- Wu, W. Q., Hou, X. M., Li, M., Dou, S. X., and Xi, X. G. (2015). BLM unfolds G-quadruplexes in different structural environments through different mechanisms. *Nucleic Acids Res.* 43, 4614–4626. doi: 10.1093/nar/gkv361
- Xi, X. G., and Deprez, E. (2010). Monitoring helicase-catalyzed DNA unwinding by fluorescence anisotropy and fluorescence cross-correlation spectroscopy. *Methods* 51, 289–294. doi: 10.1016/j.ymeth.2010.02.022
- Yao, Z., Aboualazadeh, F., Kroll, J., Akula, I., Snider, J., Lyakisheva, A., et al. (2020). Split intein-mediated protein ligation for detecting protein-protein interactions and their inhibition. *Nat. Commun.* 11:2440. doi: 10.1038/s41467-020-16299-1
- Zhou, R., Zhang, J., Bochman, M. L., Zakian, V. A., and Ha, T. (2014). Periodic DNA patrolling underlies diverse functions of Pif1 on R-loops and G-rich DNA. *eLife* 3:e02190. doi: 10.7554/eLife.02190

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Teng, Jiang, Huang, Guo, Chen, Hou, Xi and Xu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.