

# Sortase-mediated fluorescent labeling of CRISPR complexes

Kaylee E. Dillard<sup>a</sup>, Jeffrey M. Schaub<sup>a</sup>, Maxwell W. Brown<sup>a</sup>,  
Fatema A. Saifuddin<sup>a</sup>, Yibei Xiao<sup>b</sup>, Erik Hernandez<sup>c</sup>,  
Samuel D. Dahlhauser<sup>c</sup>, Eric V. Anslyn<sup>c</sup>, Ailong Ke<sup>b</sup>,  
Ilya J. Finkelstein<sup>a,d,\*</sup>

<sup>a</sup>Department of Molecular Biosciences and Institute for Cellular and Molecular Biology, University of Texas at Austin, Austin, TX, United States

<sup>b</sup>Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY, United States

<sup>c</sup>Department of Chemistry, University of Texas at Austin, Austin, TX, United States

<sup>d</sup>Center for Systems and Synthetic Biology, University of Texas at Austin, Austin, TX, United States

\*Corresponding author: e-mail address: ifinkelstein@cm.utexas.edu

## Contents

1. Introduction	2
2. Methods	5
2.1 Overview	5
2.2 Buffers	6
2.3 Sortase purification	7
2.4 Purification of <i>Tfu</i> Cascade complexes and subunits for N-terminal sortase labeling	7
2.5 Purification of <i>Tfu</i> Cascade complexes and subunits for C-terminal sortase labeling	12
2.6 Optimization of sortase-mediated fluorescent labeling	13
3. Applications	13
3.1 Single-molecule imaging of fluorescent Cas1–Cas2, Cascade, and Cas3 on DNA curtains	13
4. Notes	15
Acknowledgments	15
Funding	15
References	15

## Abstract

Fluorescent labeling of proteins is a critical requirement for single-molecule imaging studies. Many protein labeling strategies require harsh conditions or large epitopes that can inactivate the target protein, either by decreasing the protein's enzymatic activity or by blocking protein–protein interactions. Here, we provide a detailed protocol to efficiently label CRISPR–Cas complexes with a small fluorescent peptide via sortase-mediated transpeptidation. The sortase tag consists of just a few amino acids

that are specifically recognized at either the N- or the C-terminus, making this strategy advantageous when the protein is part of a larger complex. Sortase is active at high ionic strength, 4°C, and with a broad range of organic fluorophores. We discuss the design, optimization, and single-molecule fluorescent imaging of CRISPR–Cas complexes on DNA curtains. Sortase-mediated transpeptidation is a versatile addition to the protein labeling toolkit.



## 1. Introduction

Single-molecule methodologies have provided mechanistic insights into many cellular processes. Fluorescent imaging of proteins and their substrates permits visualization of transient intermediates, such as protein conformational changes, movement on DNA substrates, and protein–protein interactions. For example, single-molecule Förster resonance energy transfer (smFRET) studies have established that Cas9 undergoes conformational rearrangements in the HNH domain during target binding and cleavage (Chen et al., 2017; Dagdas, Chen, Sternberg, Doudna, & Yildiz, 2017; Sternberg, LaFrance, Kaplan, & Doudna, 2015; Yang et al., 2018). These single-molecule studies provided novel insights into how Cas9 distinguishes between on-target DNA sites and off-target sites that are partially complementary to the gRNA. Additional single-molecule studies have elucidated the mechanism of target recognition and translocation on DNA substrates for a variety of CRISPR–Cas proteins, including Cascade, Cas3, Cas1–Cas2, Cas9, and Cas12a (Dillard et al., 2018; Jung et al., 2017; Redding et al., 2015; Singh, Mallon, et al., 2018; Singh, Wang, et al., 2018; Sternberg, Redding, Jinek, Greene, & Doudna, 2014; Xue, Zhu, Zhang, Shin, & Sashital, 2017). One unifying feature of this diverse set of assays is that the protein(s) of interest needs to be fluorescently labeled *in vitro* while retaining enzymatic functions and complex-specific interaction surfaces.

*In vitro* fluorescent labeling strategies can be subdivided into three broad categories. The first category covalently attaches a fluorescent dye to cysteines, amines, and other reactive amino acid side chains. For example, cysteines can be labeled via a maleimide coupling reaction. Maleimide conjugation to the cysteine group requires all additional surface-exposed cysteines to be removed from the protein; prior knowledge of the protein structure is essential when a large set of reactive sites are possible. The protein of interest (POI) is then reduced by a reducing agent, and the reducing agent is removed prior to incubation with the maleimide-conjugated fluorescent dye (Kim et al., 2008).

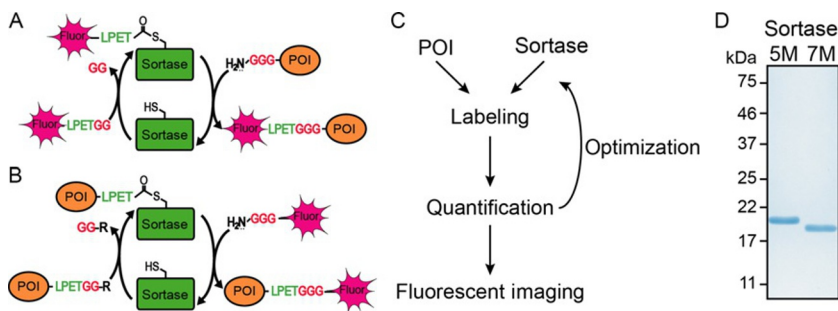
This approach is especially suitable for small proteins as all reactive cysteines have to be mutated to serines. An illustrative example is the aforementioned set of smFRET studies probing the conformational dynamics of Cas9. For that work, the authors had to replace the two native cysteines with serines and reintroduce two additional cysteines at positions that were within the expected Cy3–Cy5 dye pair Förster energy transfer distance (Sternberg et al., 2015; Yang et al., 2018). While biochemical confirmation of any fluorescent labeling scheme is always required, special care must be taken when a large set of residues need to be mutated with this labeling strategy.

Second, genetically encoded fluorescent proteins, epitope tags, or protein fragments can be fused to the POI. In this scenario, the POI contains a single fluorescent tag, which can be valuable for determining the stoichiometry or number of proteins present at a specific locus (i.e., by measuring GFP photobleaching steps). However, genetically encoded fluorescent proteins photobleach more quickly and are less bright than organic fluorophores, limiting the experimental imaging time (Cranfill et al., 2016). These epitopes are also larger than organic dyes (e.g., GFP is 27 kDa), which can block essential protein interactions or active sites (Hink et al., 2000). Protein fragments such as SNAP, CLIP, and HaloTag can be used to covalently conjugate an organic dye to the POI. This strategy significantly improves photostability of the resulting molecule compared to GFP (Cole, 2013; Gautier et al., 2008; Los et al., 2008; Sun et al., 2011). However, SNAP or CLIP adds ~29 kDa and Halo adds 33 kDa to the POI (Hansen, Rodgers, & Hoskins, 2016). This can perturb protein function and expression. Alternatively, genetically encoded epitope tags (i.e., a FLAG or HA epitope) can be added. These epitopes are then conjugated to fluorescent antibodies. This approach offers three key advantages: (1) the antibodies are commercially available; (2) multiple fluorophores can be covalently linked to each antibody for a brighter signal; and (3) the imaging strategy can be rapidly modified by using antibodies conjugated to different fluorophores. However, short epitopes can be occluded within a protein complex resulting in low labeling efficiency. In addition, the large size of a fluorescent antibody can alter the protein's activity or block potential protein binding partners. These limitations of genetically encoded fluorescent tags have prompted the development of enzymatic strategies to directly add organic fluorophores at defined positions of protein complexes.

The third broad strategy uses enzymes that covalently couple a fluorescent dye to the POI. For example, the formylglycine-generating enzyme (FGE) can be coexpressed with a protein to directly conjugate a dye to a short epitope. FGE turns the cysteine within the LCTPSR recognition

motif into formylglycine. This residue can then be coupled to a cyanine hydrazide dye (Shi et al., 2012). While this labeling strategy can reach 100% labeling efficiency, it requires a large excess of fluorophore to label the protein (e.g., >75 mM in this reported study). Using large concentrations of a fluorophore can be expensive and limited by the solubility of that particular dye. Alternatively, lower concentrations of fluorophores require a more acidic buffer and incubation at higher temperatures (e.g., pH 5.5 and 37°C) in order to get efficient labeling (Carrico, Carlson, & Bertozzi, 2007; Shi et al., 2012). These harsh conditions can cause unstable proteins to precipitate out of solution and limit the types of proteins that can be labeled via this method.

Sortase-mediated transpeptidation is a general and robust enzymatic strategy for labeling defined subunits in protein complexes. Sortase attaches proteins to the bacterial cell wall in Gram-positive bacteria. Sortase purified from *Staphylococcus aureus* uses a recognition motif of LPXTG, where X is any amino acid, to covalently link a protein containing multiple glycine residues. Recently, sortase was adapted to attach small fluorescent peptides to proteins via the N- or C-termini (Antos et al., 2009; Guimaraes et al., 2013; Theile et al., 2013). To increase the rate and efficiency of sortase labeling, two variants, 5M and 7M, were created (Fig. 1C). Both mutants have increased catalytic activity, but sortase (7M) mutant does not require addition



**Fig. 1** Overview of the sortase-mediated transpeptidation. Overview of (A) N-terminal and (B) C-terminal sortase labeling. (C) Overall labeling and optimization workflow. (D) Coomassie blue-stained SDS-PAGE gel showing purified sortase (5M) and sortase (7M). Panels (A) and (B): Schematic adapted from Guimaraes, C. P., Witte, M. D., Theile, C. S., Bozkurt, G., Kundrat, L., Blom, A. E. M., et al. (2013). Site-specific C-terminal and internal loop labeling of proteins using sortase-mediated reactions. *Nature Protocols*, 8, 1787–1799; Theile, C. S., Witte, M. D., Blom, A. E. M., Kundrat, L., Ploegh, H. L., & Guimaraes, C. P. (2013). Site-specific N-terminal labeling of proteins using sortase-mediated reactions. *Nature Protocols*, 8, 1800–1807.

of calcium to the reaction. Sortase (7M) is thus more appropriate for labeling enzymes when calcium uptake may be problematic.

Here, we describe how to optimize sortase-mediated fluorescent labeling of CRISPR–Cas complexes for single-molecule applications (Fig. 1). Key advantages are: (1) site-specific labeling of the protein via the N- or C-terminus; (2) efficient labeling in physiological buffer conditions and low temperatures; (3) minimal mutations to the POI; (4) relatively low concentrations of expensive fluorophores; and (5) attachment of a small fluorescent peptide allows for protein–protein interactions. We provide a step-by-step protocol for both N- and C-terminal labeling of *Thermobifida fusca* (Tfu) Type I-E CRISPR complexes and their subsequent reconstitution into a protein complex. In short, sortase-mediated fluorescent labeling is a broadly applicable strategy for a wide range of proteins complexes.



## 2. Methods

### 2.1 Overview

The Type I-E CRISPR system is comprised of Cascade, an 11-subunit crRNA-containing surveillance complex, the Cas1–Cas2 integrase, and Cas3 helicase/nuclease. Sortase-mediated fluorescent labeling of these sub-components allowed us to overcome multiple challenges that we and others encountered with fluorescent labeling of these proteins. For example, both Cas1–Cas2 and Cas3 directly interact with Cascade to form the primed acquisition complex (Dillard et al., 2018). Large fluorescent tags interfere with Cas1 and Cas2 activity (Redding et al., 2015), and Cas3 helicase/nuclease activity on DNA curtains (data not shown). Placement of an epitope tag on the N-terminus of Cas3 also decreases nuclease activity. Efficient labeling of these Type I-E CRISPR–Cas proteins requires a labeling strategy that works under high salt concentrations and adds a small organic fluorophore. Sortase-mediated transpeptidation allowed us to label individual CRISPR–Cas subunits.

We purified both sortase mutants (i.e., 7M and 5M), along with the POIs. Sortase labeling efficiency was optimized for each POI by testing three main parameters: the sortase mutant, incubation times, and temperatures. Optimization can be done after the POI has been isolated via a single purification column and while other proteins are present in the reaction (e.g., SUMO protease). Once optimal fluorescent labeling was determined, the reaction was scaled up and resulting labeled protein purified to remove

sortase and excess fluorescent peptide. Then, the protein's labeling efficiency was determined via a plate reader assay prior to fluorescent imaging. Using single-molecule DNA curtains, we show that these fluorescent CRISPR–Cas proteins can navigate DNA and interact with other protein partners.

## 2.2 Buffers

### 2.2.1 Sortase purification buffers

1. Sortase Lysis Buffer: 50 mM Tris–HCl (pH 7.5), 150 mM NaCl, 10 mM imidazole, 10% (v/v) glycerol, supplemented with 1 × HALT protease cocktail (ThermoFisher), and 1 mM phenylmethanesulfonyl fluoride (PMSF, Sigma–Aldrich).
2. Sortase Elution Buffer: Sortase Lysis Buffer supplemented with 500 mM imidazole.
3. Sortase Dialysis Buffer: 50 mM Tris–HCl (pH 8.0) and 150 mM NaCl.

### 2.2.2 Cas1–Cas2 purification buffers

1. Cas1–Cas2 Buffer A: 20 mM HEPES (pH 7.5), 10 mM imidazole, 500 mM NaCl.
2. Cas1–Cas2 Buffer B: 20 mM HEPES (pH 7.5), 500 mM imidazole, 500 mM NaCl.
3. Cas1–Cas2 Gel Filtration Buffer: 20 mM HEPES (pH 7.5), 500 mM NaCl, 10% glycerol.

### 2.2.3 Cascade purification buffers

1. Cascade Buffer A: 20 mM HEPES (pH 7.5), 500 mM NaCl.
2. Cascade Buffer B: 20 mM HEPES (pH 7.5), 500 mM NaCl, 5 mM desthiobiotin.
3. Cascade Gel Filtration Buffer: 10 mM Tris–HCl (pH 7.5), 150 mM NaCl, 5 mM DTT.
4. His Wash Buffer: 20 mM HEPES (pH 7.5), 500 mM NaCl, 10 mM imidazole.
5. His Elution Buffer: 20 mM HEPES (pH 7.5), 500 mM NaCl, 500 mM imidazole.

### 2.2.4 Cas3 purification buffers

1. Cas3 Buffer A: 20 mM HEPES (pH 7.5), 10 mM imidazole, 500 mM NaCl.

2. Cas3 Buffer B: 20 mM HEPES (pH 7.5), 500 mM imidazole, 500 mM NaCl.
3. Cas3 Gel Filtration Buffer: 20 mM HEPES (pH 7.5), 150 mM NaCl, 10% glycerol.

## 2.3 Sortase purification

The plasmids for sortase 7M (Addgene plasmid # 51141) and sortase 5M (Addgene plasmid # 51140) were purchased from Addgene and were generous gifts from Hidde Ploegh (Shi et al., 2014).

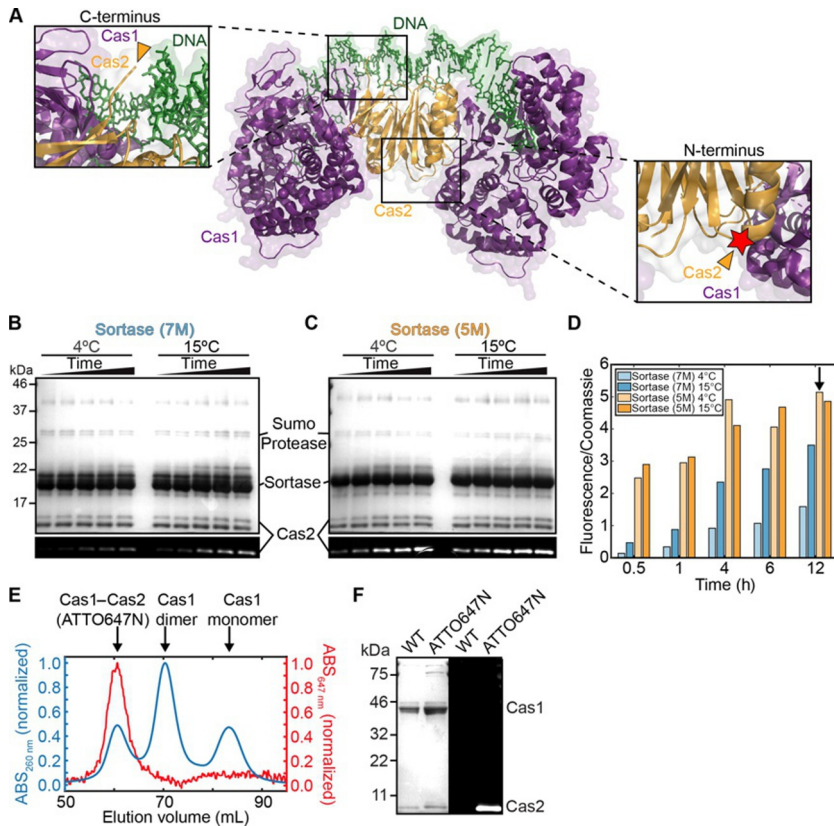
1. Transform BL21 (Novagen, 69450) *Escherichia coli* cells with either a plasmid for sortase 7M (pIF170) or 5M (pIF185).
2. Inoculate 2 L of LB with 15 mL of a starter culture and grow at 37°C to an OD<sub>600</sub> of 0.6. Induce the culture with 0.5 mM IPTG overnight at 25°C.
3. Harvest the cells by centrifuging for 15 min at 6000 × *g*.
4. Resuspend the pellets with 50 mL of Sortase Lysis Buffer.
5. Sonicate the cells to lyse and centrifuge the lysate in an ultracentrifuge at 35,000 × *g* for 45 min at 4°C.
6. Apply the supernatant onto a 1.5-mL bed volume of Ni-NTA resin (ThermoFisher) preequilibrated with Sortase Lysis Buffer.
7. Elute 1 mL fractions of sortase with Sortase Elution Buffer.
8. Pool the fractions and dialyze overnight in Sortase Dialysis Buffer.
9. Concentrate sortase to ~500 μM using a spin concentrator (EMD Millipore, UFC901008) and flash-freeze aliquots in liquid nitrogen. Store at −80°C.
10. Determine protein concentration by comparing to a BSA standard curve using SDS-PAGE.

## 2.4 Purification of *Tfu* Cascade complexes and subunits for N-terminal sortase labeling

### 2.4.1 Cas1–Cas2 purification and sortase labeling of Cas2

Fig. 2 describes the overall strategy for purifying a fluorescent Cas1–Cas2 complex. Cas2 has a GGG sortase epitope on the N-terminus after the SUMO tag. Cas2 and Cas1 are purified separately, and Cas2 is labeled with an ATTO647N dye. The Cas1–Cas2 complex is reconstituted from Cas1 and fluorescent Cas2 monomers.

1. Transform *E. coli* BL21 (DE3) cells with either a plasmid for Cas1 (pIF201) or Cas2 (pIF212).



**Fig. 2** N-terminal sortase labeling of Cas1–Cas2. (A) Structure of Cas1–Cas2 complexed with DNA (adapted from PDB: [5D55](#), [Wright et al., 2017](#)). Cas2 (orange) C-terminus is largely inaccessible. The N-terminus is positioned closer to the solvent accessible space between the Cas1 (purple) homotetramer. Orange arrows denote the N- and C-termini of Cas2 in the insets. Red star denotes approximate position of fluorescent label. We optimized N-terminal labeling of Cas2 with (B) sortase (7M) and (C) sortase (5M). Time points: 0.5, 1, 4, 6, 12 h. Top: Coomassie blue staining of an SDS-PAGE gel of the reaction; bottom: fluorescence imaging of the Cas2 band. (D) Quantification of fluorescence signal from (B) and (C). Arrow denotes conditions used for large-scale labeling. (E) Gel filtration chromatography trace showing reconstitution of the fluorescent Cas1–Cas2 complex. (F) SDS-PAGE gel of Coomassie blue (left) and fluorescence (right) of unmodified (WT) and ATTO647N-labeled Cas1–Cas2.

2. Inoculate two 1 L cultures of LB with 15 mL of a starter culture and grow at 37°C to an OD<sub>600</sub> of 0.6. Induce the culture with 0.5 mM IPTG and grow at 18°C overnight.
3. Harvest the cells by centrifuging for 15 min at 6000 × *g*.



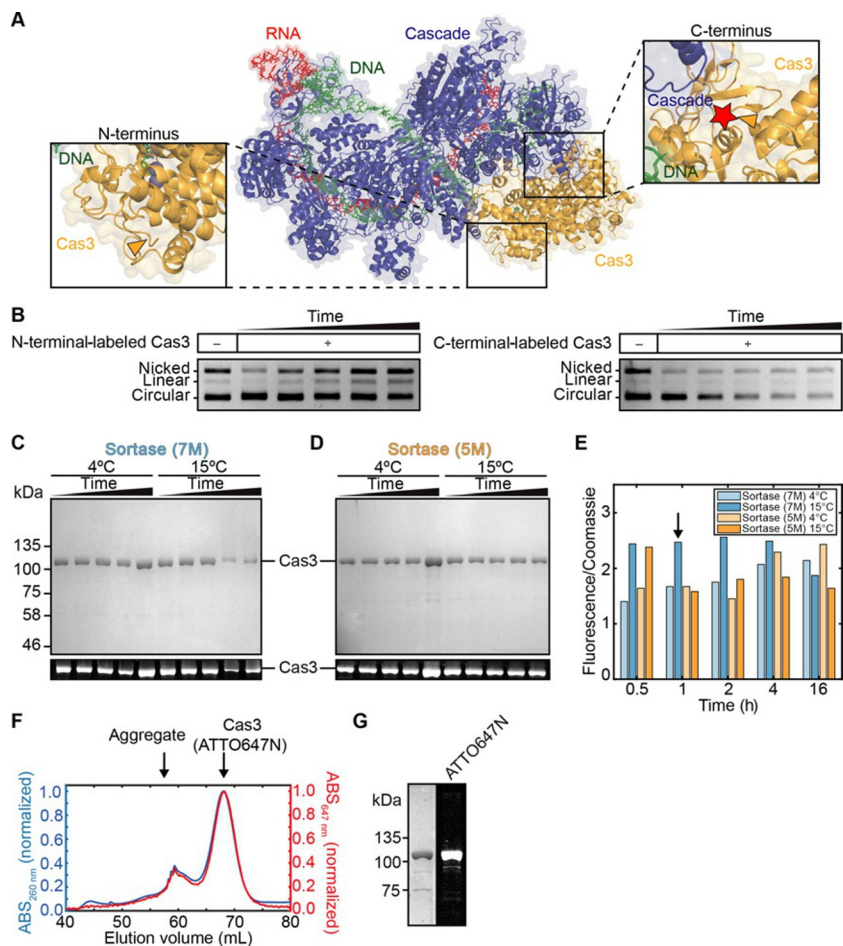
4. Resuspend the pellets with 35 mL of Cas1–Cas2 Buffer A supplemented with 10 µg/mL DNase and 1 × Halt Protease inhibitor and homogenize to lyse the cells.
5. Clarify the lysate by centrifuging in an ultracentrifuge at 35,000 × *g* for 35 min at 4°C.
6. Equilibrate two 1 mL HisTrap columns with Cas1–Cas2 Buffer A. Apply the supernatant containing either Cas1 or Cas2 to the two 1 mL HisTrap columns in tandem.
7. Wash the columns with Cas1–Cas2 Buffer A.
8. Elute either Cas1 or Cas2 with Cas1–Cas2 Buffer B. Determine the protein concentration using a Bradford assay.
9. For Cas1, incubate SUMO protease with Cas1 overnight at 4°C in a 1:50 *M* ratio with Cas1.
10. For Cas2, incubate SUMO protease and 3 × sortase 5M, 10 mM CaCl<sub>2</sub>, and 5 × LPETGG fluorescent peptide for 12 h at 4°C.
11. To remove sortase 5M from Cas2, apply the reaction to a 5-mL HisTrap column equilibrated in Cas1–Cas2 Buffer A. Collect the flow-through that contains Cas2.
12. Incubate Cas2 with 4 × Cas1 for 1 h at 4°C.
13. Spin concentrate the Cas1–Cas2 complex to 700 µL (EMD Millipore, UFC903024).
14. Isolate the Cas1–Cas2 complex on a HiPrep Sephacryl S-200 HR column pre-equilibrated in Cas1–Cas2 Gel Filtration Buffer.
15. Collect the peak fractions containing the fluorescent Cas1–Cas2 complex and flash-freeze aliquots in liquid nitrogen to be stored at –80°C.

#### 2.4.2 Cascade purification and sortase labeling of Cse1

Fig. 3 describes the overall labeling strategy for N-terminal proteins. Cse1 has a GGG N-terminal sortase epitope after the SUMO tag. Cse1 is purified and labeled separately from the rest of the Cascade complex. Sortase labeling optimization of Cse1 was carried out as above for Cas2 (see Section 2.6). The N-terminus of Cse1 was labeled because the C-terminus is occluded by other protein subunits within the complex (Fig. 3A). After sortase labeling, the Cascade complex is reconstituted.

##### **Cse1 purification and sortase labeling:**

1. Transform *E. coli* BL21 (DE3) cells with the Cse1 plasmid (pIF181).
2. Inoculate one 2 L culture of LB with 15 mL of a starter culture and grow at 37°C to an OD<sub>600</sub> of 0.8.



**Fig. 3** C-terminal sortase labeling of Cas3. (A) Structure of Cascade–Cas3 complexed with DNA (adapted from PDB: 6C66, Xiao, Luo, Dolan, Liao, & Ke, 2018). Cas3 (orange) N- and C-termini are relatively solvent exposed away from DNA (green) and Cascade (blue). Orange arrows denote the N- and C-termini of Cas3 in the insets. Red star denotes approximate position of fluorescent label. (B) Plasmid digests show that C-terminally labeled Cas3 is more active than N-terminally labeled protein. Time points of 10, 30, 60, 90, 120 min. (C) Sortase (7M) and (D) sortase (5M) were evaluated for C-terminal labeling of Cas3. Time points: 0.5, 1, 2, 4, 16 h. SDS-PAGE gel of Coomassie blue (top) and fluorescence (bottom) of the indicated proteins. (E) Quantification of fluorescent signals from (C) and (D). Arrow denotes conditions used for large-scale labeling. (F) Gel filtration chromatography trace showing that the fluorescent signal coelutes with Cas3. (G) SDS-PAGE gel of fluorescent Cas3. Left: Coomassie blue; right: fluorescent detection.

3. Induce with 1 mM IPTG and grow at 25°C overnight.
4. Harvest the cells by centrifuging for 15 min at 6000 × *g*.
5. Resuspend the pellets with 35 mL of Cascade Buffer A supplemented with 10 µg/mL DNase and 1 × Halt Protease inhibitor. Lyse the cells with sonication.
6. Clarify the lysate by centrifuging in an ultracentrifuge at 35,000 × *g* for 35 min at 4°C.
7. Equilibrate a 5-mL bed volume Strep-Tactin Superflow 50% suspension resin column with Cascade Buffer A and apply the supernatant.
8. Rinse the column with 100 mL of Cascade Buffer A.
9. Elute with 20 mL of Cascade Buffer B.
10. Spin concentrate Cse1 to 1 mL.
11. Incubate Cse1 with 2 × SUMO protease at 4°C for 12 h.
12. Add 3 × sortase 5M, 10 mM CaCl<sub>2</sub>, and 5 × LPETGG fluorescent peptide to Cse1 for 1 h at 37°C.
13. Separate Cse1 from sortase via a HiPrep Sephacryl S-200 HR column (GE) in Cascade Buffer A.

**Purification of the Cascade complex:**

1. Transform *E. coli* BL21 (DE3) cells with the Cascade B-E plasmid (pIF164) and crRNA (pIF165).
2. Inoculate one 2 L culture of LB with 15 mL of a starter culture and grow at 37°C to an OD<sub>600</sub> of 0.8.
3. Induce with 1 mM IPTG and grow at 25°C overnight.
4. Harvest the cells by centrifuging for 15 min at 6000 × *g*.
5. Resuspend the pellets with 35 mL of Cascade Buffer A supplemented with 10 µg/mL DNase and 1 × Halt Protease inhibitor. Lyse the cells with sonication.
6. Clarify the lysate by centrifuging in an ultracentrifuge at 35,000 × *g* for 35 min at 4°C.
7. Apply the supernatant to a 2-mL bed volume Ni-NTA column preequilibrated in Cascade Buffer A.
8. Wash with 100 mL His Wash Buffer.
9. Elute with 6 mL His Elution Buffer.
10. Spin concentrate to 1 mL and dialyze overnight into Cascade Buffer A.

**Reconstitution of a fluorescent Cascade complex:**

1. Combine fluorescent Cse1 and the rest of the Cascade complex in a 1:1 ratio.
2. Reconstitute the complex via a step-down NaCl dialysis.

- a. Using SnakeSkin dialysis tubing (cat# 68100), incubate in Cascade Buffer A containing 500 mM NaCl for 3 h.
- b. Dialyze at 300 mM NaCl for 3 h.
- c. Finally, dialyze at 150 mM NaCl for 3 h.
3. Isolate the fluorescent Cascade complex via a HiPrep Sephacryl S-200 HR column (GE) in Cascade Gel Filtration Buffer.
4. Collect the peak fractions and flash-freeze in liquid nitrogen to be stored at  $-80^{\circ}\text{C}$ .

## 2.5 Purification of Tfu Cascade complexes and subunits for C-terminal sortase labeling

### 2.5.1 Cas3 purification and sortase labeling

Cas3 contains a C-terminal LPETGG-TwinStrep motif for sortase labeling. The sortase reaction was optimized as described earlier for Cas2. After fluorescent labeling, the protein is sized on a gel filtration column to remove free dye and sortase protein.

1. Transform *E. coli* BL21 (DE3) cells with the Cas3 plasmid (pIF218).
2. Inoculate one 2 L culture of LB with 15 mL of a starter culture and grow at  $37^{\circ}\text{C}$  to an OD<sub>600</sub> of 0.6.
3. Induce with 0.5 mM IPTG and grow at  $18^{\circ}\text{C}$  overnight.
4. Harvest the cells by centrifuging for 15 min at  $6000 \times g$ .
5. Resuspend the pellets with 35 mL of Cas3 Buffer A supplemented with 10  $\mu\text{g}/\text{mL}$  DNase and  $1 \times$  Halt Protease inhibitor. Homogenize the cells.
6. Clarify the lysate by centrifugation at 15,000 rpm for 60 min at  $4^{\circ}\text{C}$ .
7. Equilibrate two 1 mL HisTrap columns in tandem with Cas3 Buffer A and apply the supernatant.
8. Elute Cas3 and measure the protein concentration of the 1 mL peak fraction via a Bradford assay. If the protein concentration is below 20  $\mu\text{M}$ , spin concentrate.
9. Add  $5 \times$  GGG fluorescent peptide and  $5 \times$  sortase 7M relative to the Cas3 concentration.
10. Incubate the reaction at  $15^{\circ}\text{C}$  for 1 h.
11. Isolate fluorescent Cas3 by applying to a HiPrep Sephacryl S-200 HR column (GE) equilibrated in Cas3 Gel Filtration Buffer.
12. Collect the fractions containing fluorescent Cas3 and spin concentrate to about 1 mL (EMD Millipore, UFC903024).
13. Aliquot and flash-freeze the protein in liquid nitrogen before storing at  $-80^{\circ}\text{C}$ .

## 2.6 Optimization of sortase-mediated fluorescent labeling

Sortase labeling of proteins requires optimization of three variables: sortase variants, temperature, and time.

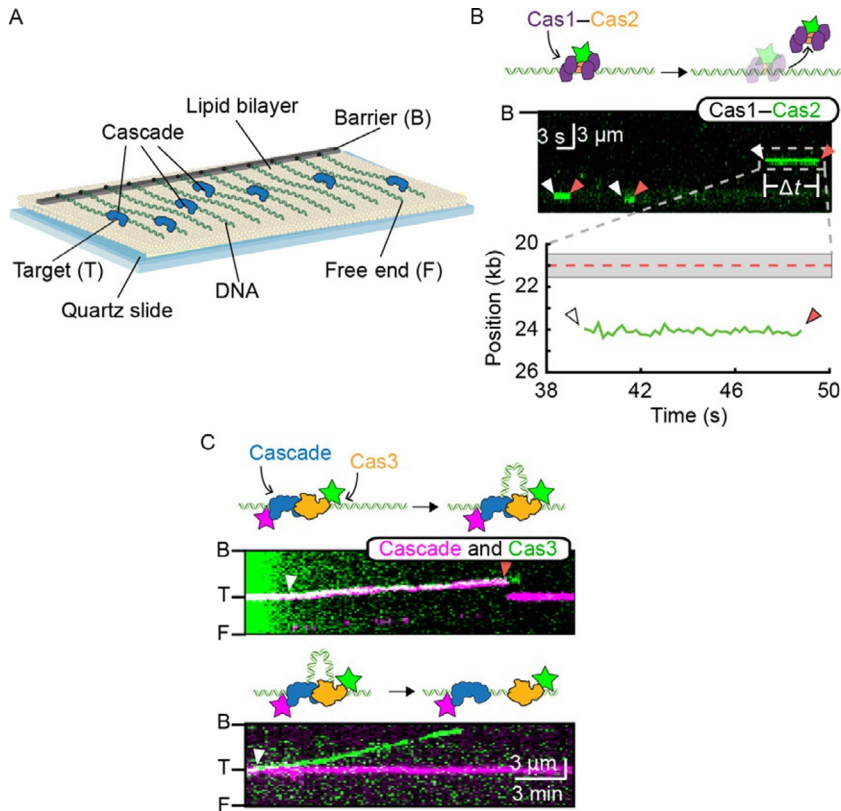
1. Prepare two master mixes with a total volume of 160  $\mu\text{L}$  containing either the sortase (5M) or sortase (7M). Each master mix includes:
  - a. Protein to be labeled (1–20  $\mu\text{M}$  is recommended).
  - b. 3–5  $\times$  sortase compared to the concentration of the protein to be labeled.
  - c. 5  $\times$  fluorescent peptide compared the concentration of the protein to be labeled.
  - d. 10 mM  $\text{CaCl}_2$  for the master mix containing the sortase 5M variant.
2. Divide each master mix into 15 separate aliquots of 10  $\mu\text{L}$  and incubate at the specific time and temperature.
  - a. Test sortase labeling at 4°C, 15°C, and 37°C. Also, test different incubation times (i.e., 0.5, 1, 4, 6, and 12 h).
3. Quench the reaction by flash freezing the aliquot in liquid nitrogen and storing at –20°C until all reactions are complete.
4. Run all reactions on an SDS-PAGE gel and acquire a fluorescent image of the gel using a Typhoon.
5. Stain the gel with Coomassie to determine the protein concentration.
6. Calculate the ratio of fluorescence intensity in the Typhoon image to protein concentration in the Coomassie-stained image for each condition to determine the best labeling condition.



## 3. Applications

### 3.1 Single-molecule imaging of fluorescent Cas1–Cas2, Cascade, and Cas3 on DNA curtains

DNA curtains were assembled as described previously (Gallardo et al., 2015; Schaub, Zhang, Soniat, & Finkelstein, 2018; Soniat et al., 2017). Briefly, biotinylated lipids were deposited on a microfluidic device containing chromium barriers. DNA was then tethered to the lipid bilayer via a streptavidin-biotin linkage. With buffer flow, the DNA is extended (Fig. 4A). Due to the difficulty of labeling Cas1–Cas2, little was known about how this complex transiently binds nonspecific DNA. To test this, we added fluorescent Cas1–Cas2 to double-tethered DNA curtains. Cas1–Cas2 transiently sampled the DNA nonspecifically via 3D collisions with a short lifetime (Fig. 4B).



**Fig. 4** Single-molecule imaging of fluorescent Cascade, Cas1–Cas2, and Cas3 on DNA curtains. (A) Schematic of DNA curtains. Lipid-tethered DNA molecules are organized at chromium diffusion barriers. Continuous buffer flow keeps the DNA extended for single-molecule imaging. (B) Cas1–Cas2 lifetime on DNA. Cartoon (*top*), kymograph (*middle*), and analysis (*bottom*) of fluorescent Cas1–Cas2 binding and dissociating from DNA. DNA binding is independent of target recognition (*red dotted line*) (C) Cartoon and kymograph of Cas3 (*green*) complex formation with Cascade (*magenta*). Fluorescent Cas3 is visualized as translocating with (*top*) or separately (*bottom*) from Cascade indicating two modes of translocation.

During interference, Cascade and Cas3 degrade single-stranded foreign DNA. Cas3 is a helicase/nuclease, so we fluorescently labeled Cas3 via a sortase-mediated reaction to see this translocation on DNA. We found Cas3 moves on DNA via two main modes: (1) remaining associated with Cascade as a complex and (2) independently away from Cascade at the target sequence (Fig. 4C). Thus, sortase-mediated labeling of the individual CRISPR–Cas proteins provides a flexible strategy for labeling challenging protein complexes.



## 4. Notes

1. We found that the extent of C-terminal sortase labeling is significantly improved when the epitope is extended past the minimal recognition site (LPETG). In our studies, the C-terminus of the POI had an LPETGG-TwinStrep epitope to remove unlabeled complexes.
2. The sortase (5M) reaction includes 10 mM CaCl<sub>2</sub>. Use sortase (7M) when labeling proteins that require a metal cofactor as this avoids uptake of the Ca<sup>2+</sup> metal ion by the POI. We still recommend testing both variants whenever possible because labeling efficiency varies for each POI (Figs. 2D and 3E).
3. We noted that N-terminal sortase labeling can sometimes also non-specifically label other proteins in a complex. Care should be taken to verify that labeling is specific to a single subunit in a complex.
4. Cas1 and Cas2 precipitate out of solution when the NaCl concentration is below 500 mM. When SUMO protease and/or sortase were added to the reaction, the buffer was adjusted to maintain the 500 mM NaCl concentration.

## Acknowledgments

We are very grateful to Dr. Hidde Ploegh for providing sortase plasmids via Addgene.

## Funding

This work was supported by the National Institutes of Health (GM097177 and GM120554 to I.J.F. and F31GM125201 to K.E.D.) and the Welch Foundation (F-I808 to I.J.F.).

## References

- Antos, J. M., Chew, G.-L., Guimaraes, C. P., Yoder, N. C., Grotenbreg, G. M., Popp, M. W.-L., et al. (2009). Site-specific N- and C-terminal labeling of a single polypeptide using sortases of different specificity. *Journal of the American Chemical Society*, 131, 10800–10801.
- Carrico, I. S., Carlson, B. L., & Bertozzi, C. R. (2007). Introducing genetically encoded aldehydes into proteins. *Nature Chemical Biology*, 3, 321–322.
- Chen, J. S., Dagdas, Y. S., Kleinstiver, B. P., Welch, M. M., Sousa, A. A., Harrington, L. B., et al. (2017). Enhanced proofreading governs CRISPR–Cas9 targeting accuracy. *Nature*, 550, 407–410.
- Cole, N. B. (2013). Site-specific protein labeling with SNAP-tags. *Current Protocols in Protein Science*, 73, 30.1.1–30.1.16.
- Cranfill, P. J., Sell, B. R., Baird, M. A., Allen, J. R., Lavagnino, Z., de Gruiter, H. M., et al. (2016). Quantitative assessment of fluorescent proteins. *Nature Methods*, 13, 557–562.



- Dagdas, Y. S., Chen, J. S., Sternberg, S. H., Doudna, J. A., & Yildiz, A. (2017). A conformational checkpoint between DNA binding and cleavage by CRISPR-Cas9. *Science Advances*, 3, eaao0027.
- Dillard, K. E., Brown, M. W., Johnson, N. V., Xiao, Y., Dolan, A., Hernandez, E., et al. (2018). Assembly and translocation of a CRISPR-Cas primed acquisition complex. *Cell*, 175, 934–946, e15.
- Gallardo, I. F., Pasupathy, P., Brown, M., Manhart, C. M., Neikirk, D. P., Alani, E., et al. (2015). High-throughput universal DNA curtain arrays for single-molecule fluorescence imaging. *Langmuir*, 31, 10310–10317.
- Gautier, A., Juillerat, A., Heinis, C., Corrêa, I. R., Kindermann, M., Beaufils, F., et al. (2008). An engineered protein tag for multiprotein labeling in living cells. *Chemistry & Biology*, 15, 128–136.
- Guimaraes, C. P., Witte, M. D., Theile, C. S., Bozkurt, G., Kundrat, L., Blom, A. E. M., et al. (2013). Site-specific C-terminal and internal loop labeling of proteins using sortase-mediated reactions. *Nature Protocols*, 8, 1787–1799.
- Hansen, S. R., Rodgers, M. L., & Hoskins, A. A. (2016). Fluorescent labeling of proteins in whole cell extracts for single-molecule imaging. *Methods in Enzymology*, 581, 83–104.
- Hink, M. A., Griep, R. A., Borst, J. W., van Hoek, A., Eppink, M. H. M., Schots, A., et al. (2000). Structural dynamics of green fluorescent protein alone and fused with a single chain Fv protein. *Journal of Biological Chemistry*, 275, 17556–17560.
- Jung, C., Hawkins, J. A., Jones, S. K., Xiao, Y., Rybarski, J. R., Dillard, K. E., et al. (2017). Massively parallel biophysical analysis of CRISPR-Cas complexes on next generation sequencing chips. *Cell*, 170, 35–47.e13.
- Kim, Y., Ho, S. O., Gassman, N. R., Korlann, Y., Landorf, E. V., Collart, F. R., et al. (2008). Efficient site-specific labeling of proteins via cysteines. *Bioconjugate Chemistry*, 19, 786–791.
- Los, G. V., Encell, L. P., McDougall, M. G., Hartzell, D. D., Karassina, N., Zimprich, C., et al. (2008). HaloTag: A novel protein labeling technology for cell imaging and protein analysis. *ACS Chemical Biology*, 3, 373–382.
- Redding, S., Sternberg, S. H., Marshall, M., Gibb, B., Bhat, P., Guegler, C. K., et al. (2015). Surveillance and processing of foreign DNA by the Escherichia coli CRISPR-Cas system. *Cell*, 163, 854–865.
- Schaub, J. M., Zhang, H., Soniat, M. M., & Finkelstein, I. J. (2018). Assessing protein dynamics on low-complexity single-stranded DNA curtains. *Langmuir*.
- Shi, X., Jung, Y., Lin, L.-J., Liu, C., Wu, C., Cann, I. K. O., et al. (2012). Quantitative fluorescent labeling of aldehyde-tagged proteins for single-molecule imaging. *Nature Methods*, 9, 499–503.
- Shi, J., Kundrat, L., Pishesha, N., Bilate, A., Theile, C., Maruyama, T., et al. (2014). Engineered red blood cells as carriers for systemic delivery of a wide array of functional probes. *Proceedings of the National Academy of Sciences of the United States of America*, 111, 10131–10136.
- Singh, D., Mallon, J., Poddar, A., Wang, Y., Tippana, R., Yang, O., et al. (2018). Real-time observation of DNA target interrogation and product release by the RNA-guided endonuclease CRISPR Cpf1 (Cas12a). *Proceedings of the National Academy of Sciences of the United States of America*, 115, 5444–5449.
- Singh, D., Wang, Y., Mallon, J., Yang, O., Fei, J., Poddar, A., et al. (2018). Mechanisms of improved specificity of engineered Cas9s revealed by single-molecule FRET analysis. *Nature Structural & Molecular Biology*, 25, 347–354.
- Soniat, M. M., Myler, L. R., Schaub, J. M., Kim, Y., Gallardo, I. F., & Finkelstein, I. J. (2017). Next-generation DNA curtains for single-molecule studies of homologous recombination. *Methods in Enzymology*, 592, 259–281.



- Sternberg, S. H., LaFrance, B., Kaplan, M., & Doudna, J. A. (2015). Conformational control of DNA target cleavage by CRISPR-Cas9. *Nature*, 527, 110–113.
- Sternberg, S. H., Redding, S., Jinek, M., Greene, E. C., & Doudna, J. A. (2014). DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature*, 507, 62–67.
- Sun, X., Zhang, A., Baker, B., Sun, L., Howard, A., Buswell, J., et al. (2011). Development of SNAP-tag fluorogenic probes for wash-free fluorescence imaging. *ChemBiochem*, 12, 2217–2226.
- Theile, C. S., Witte, M. D., Blom, A. E. M., Kundrat, L., Ploegh, H. L., & Guimaraes, C. P. (2013). Site-specific N-terminal labeling of proteins using sortase-mediated reactions. *Nature Protocols*, 8, 1800–1807.
- Wright, A. V., Liu, J.-J., Knott, G. J., Doxzen, K. W., Nogales, E., & Doudna, J. A. (2017). Structures of the CRISPR genome integration complex. *Science*, 357, eaao0679.
- Xiao, Y., Luo, M., Dolan, A. E., Liao, M., & Ke, A. (2018). Structure basis for RNA-guided DNA degradation by Cascade and Cas3. *Science*, 361, eaat0839.
- Xue, C., Zhu, Y., Zhang, X., Shin, Y.-K., & Sashital, D. G. (2017). Real-time observation of target search by the CRISPR surveillance complex Cascade. *Cell Reports*, 21, 3717–3727.
- Yang, M., Peng, S., Sun, R., Lin, J., Wang, N., & Chen, C. (2018). The conformational dynamics of Cas9 governing DNA cleavage are revealed by single-molecule FRET. *Cell Reports*, 22, 372–382.