


**Chemical Stepping** **Hot Paper**

 How to cite: *Angew. Chem. Int. Ed.* **2023**, *62*, e202300890

International Edition: doi.org/10.1002/anie.202300890

German Edition: doi.org/10.1002/ange.202300890

# Mobile Molecules: Reactivity Profiling Guides Faster Movement on a Cysteine Track

 Zonghua Bo<sup>+</sup>, Zhong Hui Lim<sup>+</sup>, Fernanda Duarte,<sup>\*</sup> Hagan Bayley,<sup>\*</sup> and Yujia Qing<sup>\*</sup>

**Abstract:** We previously reported a molecular hopper, which makes sub-nanometer steps by thiol-disulfide interchange along a track with cysteine footholds within a protein nanopore. Here we optimize the hopping rate (ca.  $0.1 \text{ s}^{-1}$  in the previous work) with a view towards rapid enzymeless biopolymer characterization during translocation within nanopores. We first took a single-molecule approach to obtain the reactivity profiles of individual footholds. The  $pK_a$  values of cysteine thiols within a pore ranged from 9.17 to 9.85, and the pH-independent rate constants of the thiolates with a small-molecule disulfide varied by up to 20-fold. Through site-specific mutagenesis and a pH increase from 8.5 to 9.5, the overall hopping rate of a DNA cargo along a five-cysteine track was accelerated 4-fold, and the rate-limiting step 21-fold.

## Introduction

In cells, motor proteins, such as kinesins, dyneins, and myosins, move along microtubule and actin tracks.<sup>[1]</sup> In contrast to the prevalence of protein-based tracks in nature, synthetic mobile molecules tread mostly on DNA<sup>[2–4]</sup> and small-molecule tracks,<sup>[5–7]</sup> which are often designed for ease of synthesis. Recently, we constructed small-molecule walkers and hoppers that move along a protein track comprising cysteine footholds positioned approximately 6.9 Å apart along a  $\beta$  strand within an  $\alpha$ -hemolysin ( $\alpha$ HL) pore (Figure 1a).<sup>[8,9]</sup> Cysteine is a desirable building block for a protein track, because the side-chain thiol, when deprotonated, is a powerful nucleophile. Thiolates can participate in various rapid and often reversible chemistries (e.g., thiol-disulfide interchange and arsenic(III)-thiol chemistry) that can be used to produce stepwise molecular

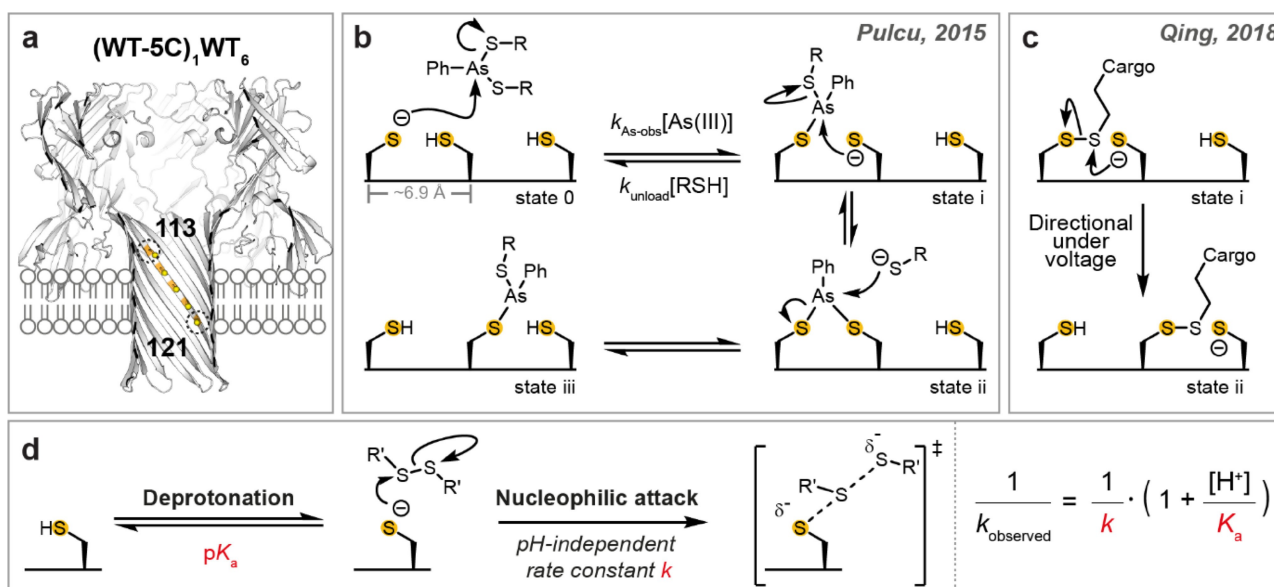
motion.<sup>[10,11]</sup> We first reported the movement of a two-legged molecular walker along a five-cysteine track within the  $\alpha$ HL pore (footholds at positions 113, 115, 117, 119 and 121) based on the substitution reactions of arsenic(III) compounds with thiols (Figure 1b).<sup>[8]</sup> Later, we developed a single-legged molecular hopper that carries a DNA cargo and moves along the cysteine track using thiol-disulfide interchange (Figure 1c).<sup>[9]</sup> Under an electric potential, the hopping motion was autonomous, processive, and directional. The molecular hopper moves at a stepping rate averaging  $0.13 \text{ s}^{-1}$  at pH 8.5. The rate varies by up to an order of magnitude depending on the position on the track.<sup>[9]</sup> Later, the molecular hopper was elaborated into an enzymeless system to effect directional stepwise translocation of biopolymers for sequence analysis.<sup>[12]</sup> Future practical use of the enzymeless system will require both a longer track and more rapid stepping. To address the latter, we assessed the reactivity of individual cysteine footholds comprising the molecular track to determine how the kinetics of mobile molecules on cysteine tracks can be manipulated.

We first took a nanoreactor approach to probe the reactivity of individual cysteine footholds at the single-molecule level.<sup>[13]</sup> Reactivity in the context of the thiol-disulfide interchange that underlies chemical hopping was determined from the  $pK_a$  values of the cysteine thiols and the pH-independent rate constants for nucleophilic attack by the corresponding thiolates on a set electrophile (Figure 1d). Single-cysteine mutations were made at positions 113, 115, 117, 119, 121 or 123. Nanoreactors containing one mutant subunit were prepared, and named (M113C)<sub>1</sub>WT<sub>6</sub>, (T115C)<sub>1</sub>WT<sub>6</sub>, (T117C)<sub>1</sub>WT<sub>6</sub>, (G119C)<sub>1</sub>WT<sub>6</sub>, (N121C)<sub>1</sub>WT<sub>6</sub>, and (N123C)<sub>1</sub>WT<sub>6</sub>. Each nanoreactor was inserted into a planar lipid bilayer for electrical recording. In the presence of small-molecule disulfide and thiol reagents, spatially separated by the lipid bilayer, thiol-disulfide interchange reactions occurred at the cysteine site in a cycle of three steps (Figure 2a).<sup>[13]</sup> In Step 1, a disulfide molecule (RSSR') entered from the *cis* side to react with the cysteine side chain (state 1) to generate an  $\alpha$ HL-SR' mixed disulfide (state 2). In Step 2, the mixed disulfide was attacked by DL-dithiothreitol (DTT) introduced from the *trans* side to generate an  $\alpha$ HL-DTT mixed disulfide (state 3), which then self-cyclized in Step 3 (Figure 2b). These thiol-disulfide interchange reactions were monitored by changes in the ionic current passing through the  $\alpha$ HL nanoreactor under an applied transmembrane electric potential; the current reflected the bond-making and bond-breaking events in real time (Figure 2c). Rate constants were obtained at different pH values from the mean lifetimes of each current state<sup>[14]</sup>

[\*] Z. Bo,<sup>+</sup> Z. H. Lim,<sup>+</sup> F. Duarte, H. Bayley, Y. Qing  
 Department of Chemistry, University of Oxford  
 Mansfield Road, Oxford, OX1 3TA (UK)  
 E-mail: fernanda.duartegonzalez@chem.ox.ac.uk  
 hagan.bayley@chem.ox.ac.uk  
 yujia.qing@chem.ox.ac.uk

[†] These authors contributed equally to this work.

© 2023 The Authors. Angewandte Chemie International Edition published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.



**Figure 1.** Cargo translocation along a protein track. a) An  $\alpha$ HL nanoreactor containing a five-cysteine track bounded by Cys-113 and Cys-121. b) The molecular walker, formed in situ by the reaction of 4-sulphophenylarsonous acid with excess 2-(2-methoxyethoxy)ethanethiol, is referred to as As(III). The walker was first loaded onto the protein track (state 0 $\rightarrow$ i); translocation of the walker (state i $\rightarrow$ iii) was mediated by two consecutive nucleophilic substitutions.<sup>[8]</sup> c) Chemical stepping of the molecular hopper along the protein track proceeded through thiol-disulfide interchange (state i $\rightarrow$ ii). The directionality of translocation was controlled by an externally applied electric potential.<sup>[9]</sup> d) Reactivity in a thiol-disulfide interchange reaction can be separated into two components: thiol deprotonation and subsequent nucleophilic attack. The reaction parameters of individual steps were obtained by using the equation shown.

and used to calculate the  $pK_a$  values of the cysteine thiols and the pH-independent rate constants of the thiolates for attack on the disulfide.

In addition, we modified the environment in the vicinity of Cys-113 on the molecular track by mutating position 111 to alter the rate of thiol-disulfide interchange at this position. A mutation that increased the rate was found, which was later incorporated into protein tracks. Finally, chemical stepping of a DNA molecule on a modified cysteine track was tested at two pH values; faster chemical stepping was observed at a higher pH value.

## Results and Discussion

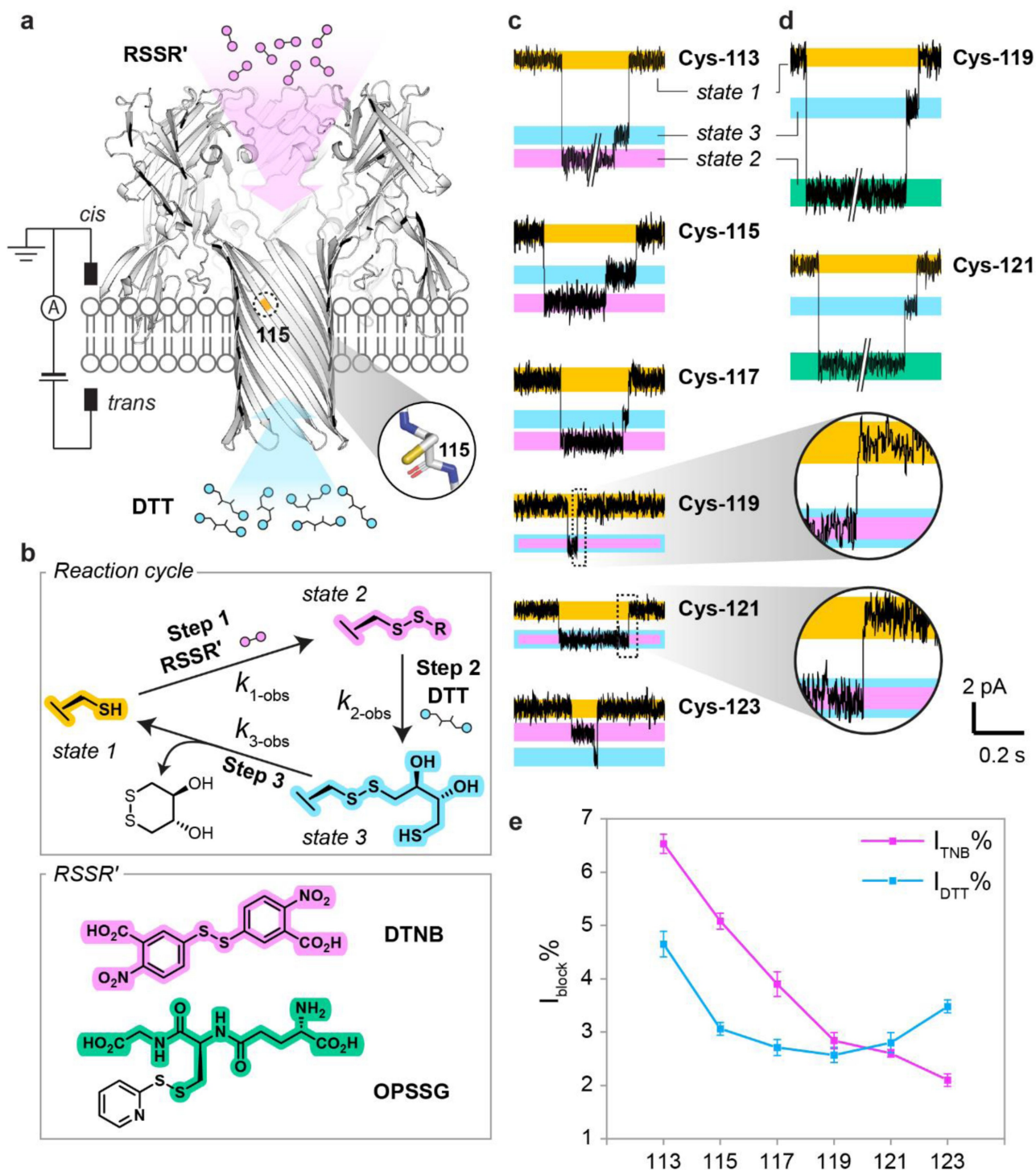
### Single-Molecule Reactivity Profiling of Individual Cysteine Residues

To monitor the three-step reaction cycle repeatedly at individual cysteine sites (i.e., Cys-113, Cys-115, Cys-117, Cys-119, Cys-121 and Cys-123) we first treated individual cysteine nanoreactors with 5,5'-dithiobis(2-nitrobenzoic acid) (DTNB) and DTT (Figure 2b). Current blockades of the mixed disulfides ( $I_{\text{TNB}}\%$  or  $I_{\text{DTT}}\%$ ) were calculated as a percentage of the open pore current (state 1). We found that the  $\alpha$ HL-TNB and  $\alpha$ HL-DTT adducts formed at positions 119 and 121 were indistinguishable in terms of current blockade ( $I_{\text{TNB}}\% = I_{\text{DTT}}\%$ , Figure 2c).  $I_{\text{TNB}}\% > I_{\text{DTT}}\%$  was found at positions 113, 115, and 117, whereas  $I_{\text{TNB}}\% < I_{\text{DTT}}\%$  was observed at position 123 (Figure 2e). Spatial blockades of the  $\alpha$ HL-TNB or  $\alpha$ HL-DTT adducts at Cys-113, Cys-117,

and Cys-121, defined as the end-to-end distance of the adduct divided by the radius of the unreacted pore at the cysteine position, were calculated over 600 ns of cumulative molecular dynamics simulations. We observed a linear correlation between the relative simulated spatial blockades and the relative experimental current blockades (see the Supporting Information Section 1 for supplementary discussion).

To resolve the current levels of  $\alpha$ HL-SR' and  $\alpha$ HL-DTT at Cys-119 or Cys-121, we treated (G119C)<sub>1</sub>WT<sub>6</sub> or (N121C)<sub>1</sub>WT<sub>6</sub> with orthopyridyl glutathionyl disulfide (OPSSG) and DTT. The  $\alpha$ HL-SG adduct blocked significantly more current than the  $\alpha$ HL-DTT adduct, allowing current level resolution (Figure 2d). However, undesirable regioselectivity was observed in Step 2.  $\alpha$ HL-SG often reacted with DTT to regenerate the nanoreactor cysteine directly without forming an  $\alpha$ HL-DTT adduct, while reaction of  $\alpha$ HL-TNB with DTT exclusively formed  $\alpha$ HL-DTT (see Supporting Information Section 2 for supplementary discussion).

Similar to the foothold thiolates involved in chemical stepping,<sup>[9]</sup> a nanoreactor cysteine thiolate acted both as the nucleophile (Step 1) and the leaving group (Step 3) during a reaction cycle. To estimate the reactivity of each cysteine thiolate, reaction cycles were recorded using DTNB and DTT at a minimum of three pH values between pH 8.0 and 9.5 using either HEPBS or AMPPO as the buffering agent (see Experimental Details in the Supporting Information). Given that the  $pK_a$  value of a cysteine residue depends on its environment within a protein,<sup>[10]</sup> an increase in the pH value during a reaction could in principle alter the thiol  $pK_a$



**Figure 2.** Single-molecule approach to the determination of  $pK_a$  values for cysteine residues. a) Reactants RSSR' and DTT, spatially separated by the lipid bilayer, were introduced from the *cis* and *trans* compartments respectively. Cys-115 is shown as an example of a residue with a reactive side chain. b) The thiol-disulfide interchange reaction cycle used for  $pK_a$  determination. In each experiment, either DTNB or OPSSG was introduced into the *cis* compartment. c) Single-channel current recordings depicting complete reaction cycles with DTNB for each single-cysteine nanoreactor. d) Reaction cycles for OPSSG. The levels in C and D were color-coded: free cysteine side chain (yellow),  $\alpha$ HL-TNB (pink),  $\alpha$ HL-SG (green) and  $\alpha$ HL-DTT (blue). For Cys-119 and Cys-121, the  $\alpha$ HL-DTT level could be discriminated from the  $\alpha$ HL-SG level, but not from the  $\alpha$ HL-TNB level (pink with blue outline). e) Percentage current blockades ( $I_{block} \%$ ) for  $\alpha$ HL-TNB and  $\alpha$ HL-DTT for positions 113 to 123.

value by 1) deprotonating neighboring ionizable amino acid side chains (if present); 2) eliciting protein conformational

changes. We expect that the  $pK_a$  values of individual cysteine thiols are minimally affected within the tested pH

range (8.0 to 9.5), because 1) the inward-facing residues within the  $\beta$  barrel either are non-ionizable or require a higher pH value for deprotonation (e.g., serine and threonine ( $pK_a > 13$ )<sup>[15,16]</sup> or lysine ( $pK_a > 10.5$ )<sup>[17,18]</sup>); 2) the  $\alpha$ HL nanopore is stable up to pH 11, with unfolding of the cap domain occurring at pH 12.<sup>[19]</sup>

The observed rate constants for Step 1 or Step 3 ( $k_{i-obs}$ ,  $i=1$  or 3) were derived from the mean lifetimes of the associated chemical states ( $\langle\tau_i\rangle$ ,  $i=1$  or 3). For the bimolecular Step 1,  $k_{1-obs} = 1/(\langle\tau_1\rangle[DTNB]_{cis})$ , while for the unimolecular Step 3,  $k_{3-obs} = 1/\langle\tau_3\rangle$ . To ensure there was no buffer catalysis, we recorded the reaction cycles in (T115C)<sub>1</sub>WT<sub>6</sub> at pH 8.8, buffered by both HEPBS and AMPSO, and observed no significant difference in  $k_{3-obs}$  obtained with the two buffer systems (see Supporting Information Section 3 for supplementary discussion).

Using the Henderson–Hasselbalch equation [Eq. (1a)], the pH-independent rate constants of the thiolate ( $k_i$ ,  $i=1$  or 3) and the  $pK_a$  values of the respective thiols for each nanoreactor were determined by fitting  $1/k_{i-obs}$  on  $[H^+]$  using linear regression [Eq. (1b)].<sup>[20]</sup>

$$k_{i-obs} = k_i / (1 + 10^{pK_a - pH}) \quad (i = 1, 3) \quad (1a)$$

$$(1/k_{i-obs}) = (1/k_i) \cdot [1 + ([H^+]/K_a)] \quad (i = 1, 3) \quad (1b)$$

For each nucleophilic thiol, we performed linear regression fitting of  $1/k_{i-obs}$  on  $[H^+]$  with 3-fold cross-validation to calculate the  $k_i$  and the  $pK_a$  values (see Supporting Information Section 4 for supplementary discussion). The  $pK_a$  values of the six cysteine residues ranged from 9.17 (Cys-113) to 9.85 (Cys-123; Table 1). The pH-independent rate constants of the cysteine thiolates for the reaction with DTNB ( $k_i$ ; Cys-113, Cys-115, Cys-117, Cys-119, Cys-121, or Cys-123) ranged from  $2.76 \pm 0.21 \times 10^3 \text{ M}^{-1} \text{ s}^{-1}$  to  $5.89 \pm 0.52 \times 10^4 \text{ M}^{-1} \text{ s}^{-1}$ , whereas those for intramolecular cyclization of the  $\alpha$ HL-DTT adducts ( $k_3$ ; Cys-113, Cys-115, Cys-117 or Cys-123) ranged from  $174 \pm 38 \text{ s}^{-1}$  to  $335 \pm 48 \text{ s}^{-1}$ . The errors of the cysteine  $pK_a$  values were less than 0.1, suggesting good data reproducibility despite various sources of experimental error (e.g., determinations of pH, concentration, and temperature) and pore-to-pore variation. In comparison to the ensemble pH-independent rate constants measured for glutathione (Table 1) with DTNB, our  $k_1$  values obtained from single-molecule experiments at Cys-115, Cys-117, Cys-119 and Cys-121 were similar (ca.  $10^4 \text{ M}^{-1} \text{ s}^{-1}$ ), while the  $k_1$  values at Cys-113 and Cys-123 were around 10-fold lower.

Our results suggest that the nucleophilicity of cysteine residues in single-cysteine nanoreactors are not linearly correlated with their  $pK_a$  values. Previous studies of thiol-disulfide interchange reactions between small-molecule thiols and disulfides revealed a Brønsted relation wherein the pH-independent rate constant of the thiolate correlated positively with the  $pK_a$  value of the nucleophilic thiol but negatively with the  $pK_a$  value of the leaving-group thiol.<sup>[20]</sup> From these studies, Brønsted coefficients of the nucleophilic thiolate were determined to be 0.4–0.5. In a protein environment, the expected Brønsted relation may be complicated by factors such as electrostatics, sterics, and solvation.<sup>[10,24–26]</sup>

**Table 1:** pH-independent rate constants and  $pK_a$  values of nucleophilic thiols in thiol-disulfide interchange reactions in single-molecule experiments and ensemble experiments.

Step 1: thiol-disulfide interchange between cysteine thiolate and DTNB, $k_1$ ( $10^3 \text{ M}^{-1} \text{ s}^{-1}$ ) <sup>[a]</sup>		
Nanopore		$pK_a$ (cysteine)
(M113C) <sub>1</sub> WT <sub>6</sub>	$2.76 \pm 0.21$	$9.17 \pm 0.04$
(T115C) <sub>1</sub> WT <sub>6</sub>	$30.0 \pm 1.5$	$9.37 \pm 0.03$
(T117C) <sub>1</sub> WT <sub>6</sub>	$20.0 \pm 2.2$	$9.31 \pm 0.05$
(G119C) <sub>1</sub> WT <sub>6</sub>	$56.2 \pm 9.3$	$9.41 \pm 0.07$
(N121C) <sub>1</sub> WT <sub>6</sub>	$58.9 \pm 5.2$	$9.59 \pm 0.05$
(N123C) <sub>1</sub> WT <sub>6</sub>	$6.30 \pm 0.50$	$9.85 \pm 0.04$
Ensemble experiments: thiol-disulfide interchange between thiolate and DTNB ( $10^3 \text{ M}^{-1} \text{ s}^{-1}$ )		
Nucleophile		$pK_a$ (thiol)
Small-molecule thiols <sup>[b]</sup>	$180 - 2890$ <sup>[b]</sup>	$9.33 - 9.80$ <sup>[b]</sup>
Glutathione thiols <sup>[c]</sup>	$35 - 40$ <sup>[c]</sup>	/
Step 3: self-cyclization of $\alpha$ HL-DTT, $k_3$ ( $\text{s}^{-1}$ ) <sup>[a]</sup>		
Nanopore		$pK_a$ ( $\alpha$ HL-DTT)
(M113C) <sub>1</sub> WT <sub>6</sub>	$174 \pm 38$	$9.07 \pm 0.10$
(T115C) <sub>1</sub> WT <sub>6</sub>	$335 \pm 48$	$9.91 \pm 0.07$
(T117C) <sub>1</sub> WT <sub>6</sub>	$194 \pm 12$	$9.23 \pm 0.03$
(N123C) <sub>1</sub> WT <sub>6</sub>	$188 \pm 15$	$9.29 \pm 0.04$

[a] Thiol-disulfide interchange reactions involved in Step 1 and Step 3 within each nanoreactor were conducted over a range of pH values between 8.0 and 9.5 in 2 M KCl, 50 mM HEPBS/AMPSO adjusted to the appropriate pH value, 20  $\mu$ M EDTA at  $20 \pm 1$  °C. The applied potential was +50 mV. The  $pK_a$  value of each nucleophilic thiol and the pH-independent rate constant of the thiolate  $k_i$  were calculated from the mean gradient and mean y-intercept obtained from linear regression fitting with 3-fold cross-validation of input values of  $[H^+]$  against the target values of  $1/k_{i-obs}$ . Standard deviations were calculated using error propagation from standard deviations associated with the mean gradient and mean y-intercept. [b] Ensemble pH-independent rate constants and  $pK_a$  values from the literature<sup>[21,22]</sup> for thiols: 2-mercaptoethanol, 1-thioglycerol, thioglycolic acid, and methyl 3-mercaptopropionate. [c] Ensemble pH-independent rate constants from the literature for glutathione.<sup>[23]</sup>

For example, a study of reactions between disulfides and cysteine thiols embedded in peptides revealed the importance of the net charge of disulfide containing molecules; Brønsted coefficients for the nucleophilic peptide thiolates were 0 and 0.8 for the negatively charged glutathione disulfide and the positively charged cystamine, respectively.<sup>[27]</sup> Given the small dataset and the narrow  $pK_a$  range covered in the present study, it is unrealistic to extrapolate a free-energy relationship. Nevertheless, we noted that the poorest nucleophiles corresponded to the cysteine residues with the lowest  $pK_a$  (Cys-113) and, surprisingly, the highest  $pK_a$  (Cys-123), with  $k_1$  values approximately an order of magnitude lower than for the remaining cysteine residues.

Previously, an arsenic(III)-based molecular walker was loaded onto a cysteine track by a thiol-arsenic interchange reaction between an  $\alpha$ HL cysteine and a thiol-arsenic(III) adduct (Figure 1b).<sup>[8]</sup> The observed rate constants of walker loading ( $k_{As-obs}$ ) at Cys-113, Cys-115, Cys-117, Cys-119, and Cys-121 were measured individually at pH 8.0, and exhibited a similar trend to that of  $k_{1-obs}$  (i.e., the observed rate constant for the reaction of a nanoreactor cysteine thiol with DTNB) at pH 8.0, with Cys-113 being the most slow-reacting (see Supporting Information Section 5 for supplementary discussion). Thiol-arsenic exchange reactions were shown to occur with stereochemical inversion,<sup>[28]</sup> suggesting a collinear transition state (S–As–S) at the arsenic center. This parallels the collinear transition state required for thiol-disulfide interchange.<sup>[29]</sup> Hence, the factors influencing the trends observed for the rate constants of thiol-disulfide interchange (e.g., the  $pK_a$  values of the cysteine thiols, the sterics of the protein environment) might have a parallel effect on thiol-arsenic exchange.

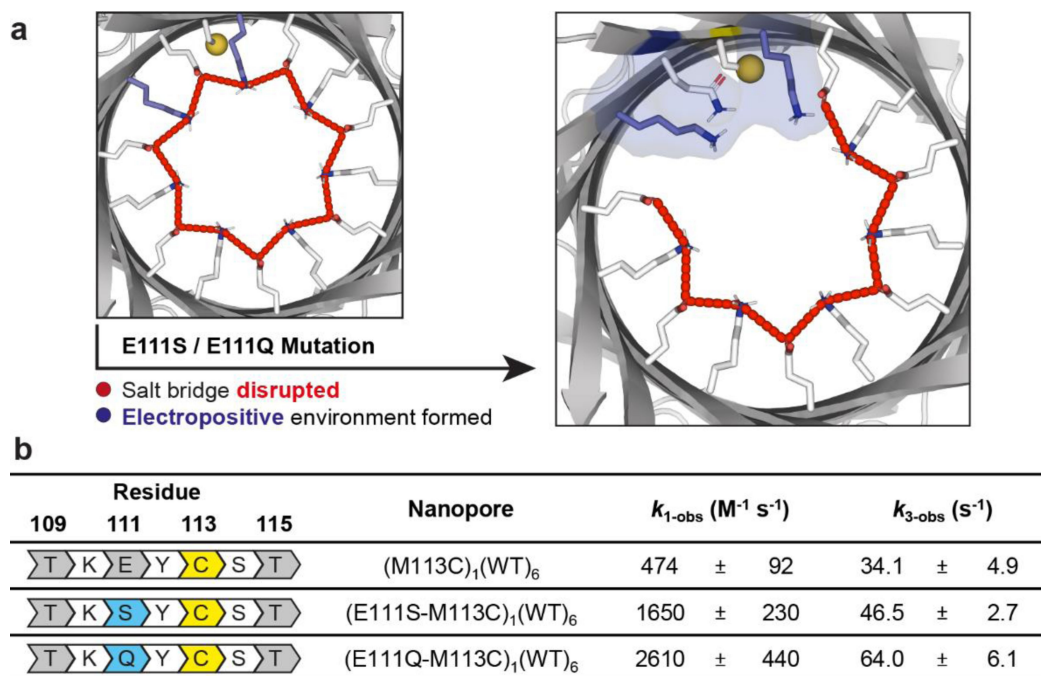
### Enhanced Cysteine Reactivity by Mutating a Neighboring Residue

Noting that the Cys-113 thiolate had the lowest  $k_1$  value, we engineered the surrounding protein environment to investigate the extent to which its reactivity could be altered. We generated two mutant single-cysteine nanoreactors, (E111S-M113C)<sub>1</sub>WT<sub>6</sub> and (E111Q-M113C)<sub>1</sub>WT<sub>6</sub>, containing Cys-113 and either a E111S or E111Q mutation in one of the seven

subunits. A net positive charge from Lys-147 is obtained when the negatively charged Glu-111 is replaced with the neutral Ser-111 or Gln-111. Consequently, the Lys-147 residues (on the mutant subunit and adjacent wild-type subunit) are no longer salt-bridged to the absent Glu-111 residue (Figure 3a).

Both (E111S-M113C)<sub>1</sub>WT<sub>6</sub> and (E111Q-M113C)<sub>1</sub>WT<sub>6</sub> gave a stable steady-state current without frequent gating under  $\pm 150$  mV, essential for the observation of chemical stepping of a DNA cargo molecule. Both mutations led to an increase in the observed rate constants of thiol-disulfide interchange (Step 1,  $k_{1-obs}$ ) between Cys-113 and DTNB by 3.5-fold (E111S) or 5.5-fold (E111Q) at pH 8.5 (Figure 3b). Similarly, the observed rate constant of DTT cyclization (Step 3,  $k_{3-obs}$ ) at pH 8.5 increased by 1.4-fold (E111S) or 1.9-fold (E111Q; Figure 3b). The electropositive local environment, contributed by the Lys-147 residues, and hydrogen-bond donation from Ser-111 or Gln-111 residues were hypothesized to stabilize the Cys-113 thiolate (i.e., lower the  $pK_a$  of Cys-113). The increased probability of thiol deprotonation increases  $k_{1-obs}$ .<sup>[31,32]</sup> The greater rate enhancements observed for the E111Q variant implied that a longer side chain enables more favorable interactions for thiol-disulfide interchange rather than steric hindrance.

The rate enhancement for DTT cyclization ( $k_{3-obs}$ ) is attributed to: a) a lower  $pK_a$  value of the  $\alpha$ HL-DTT thiol producing increased thiol deprotonation, and b) a lower  $pK_a$  value of the Cys-113 thiol resulting in the thiolate becoming a better leaving group.<sup>[20]</sup> The rate enhancement for DTT cyclization was lower than that for the preceding Step 1. We



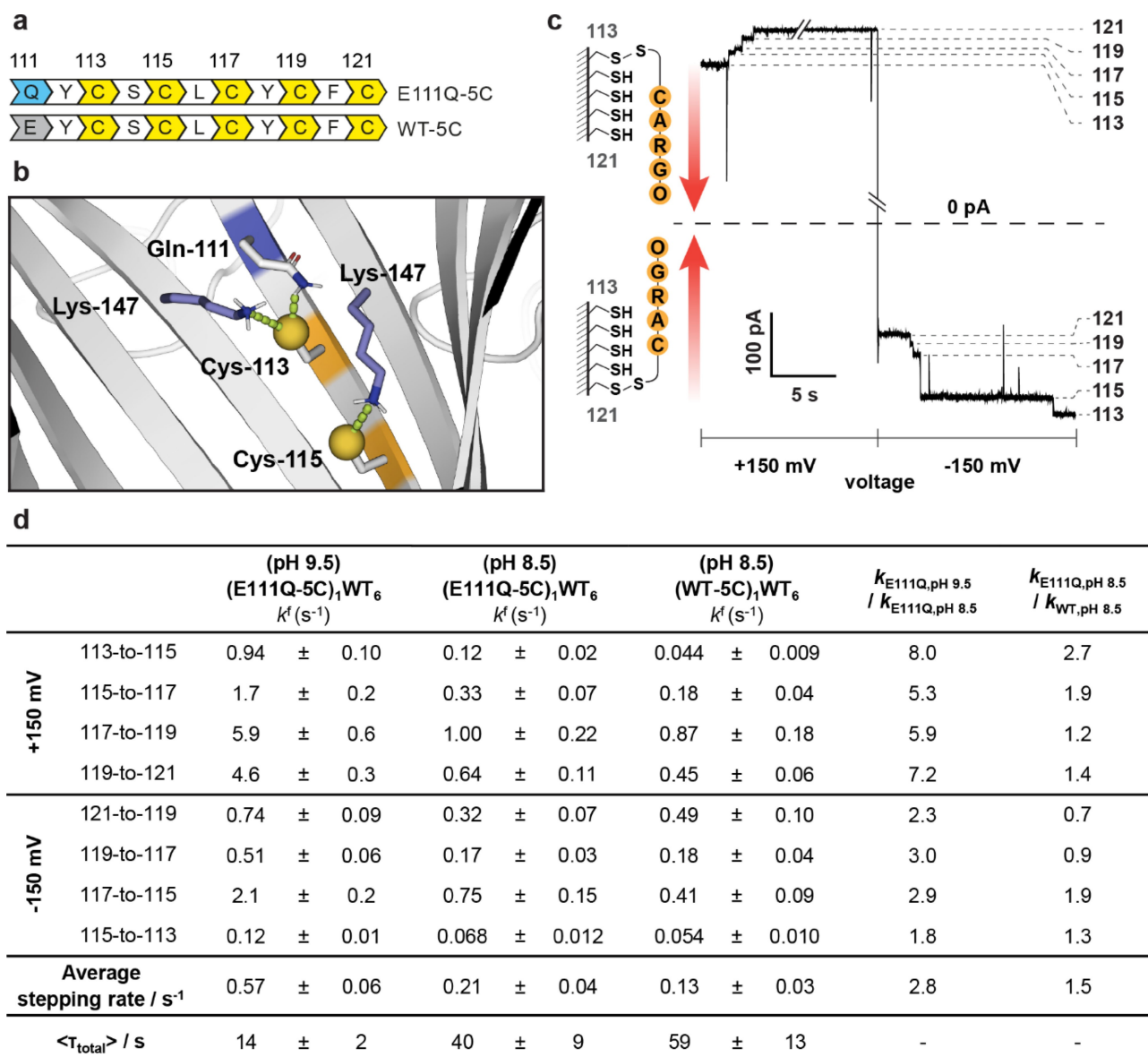
**Figure 3.** Rational redesign of an engineered track. a) Mutations E111S and E111Q are predicted to disrupt the ring of salt bridges (red dashed lines) and liberate lysine residues (colored in blue). b) Observed rate constants for nucleophilic attack by Cys-113 on DTNB ( $k_{1-obs}$ ) and intramolecular cyclization of the DTT adducts ( $k_{3-obs}$ ). Dwell-time analysis and rate constant estimations were performed by using the maximum interval likelihood algorithm of QuB software.<sup>[30]</sup> Conditions: 2 M KCl, 50 mM HEPBS, pH 8.5, 20  $\mu$ M EDTA, 20  $\pm$  1  $^{\circ}$ C.

speculate this to be due to the  $\alpha$ HL-DTT adduct extending further into the lumen of the pore, away from the Lys-147 and Ser-111 or Gln-111 residues as compared to the Cys-113 thiol.<sup>[32]</sup>

### Accelerated Chemical Stepping along Engineered Tracks

Encouraged by the ability to increase the reactivity of Cys-113, we incorporated Gln-111 into nanoreactors containing a

five-cysteine track. Using PyMOL modeling, we identified favorable hydrogen-bonding and ionic interactions within the mutant environment to stabilize the deprotonated form of either Cys-113 (N–S ca. 3.6 Å between Lys-147 and Cys-113 or N–S ca. 2.2 Å between Gln-111 and Cys-113) or Cys-115 (N–S ca. 3.3 Å between Lys-147 and Cys-115; Figure 4a). These interactions were expected to lower the  $pK_a$  values of both Cys-113 and Cys-115 thiols, promoting the formation of a nucleophilic thiolate and improving the leaving group ability of the cysteine thiolate.<sup>[20]</sup> The (E111Q-



**Figure 4.** Acceleration of hopping rates along a modified cysteine track. a) Primary sequences of the protein tracks investigated. b) Rate enhancements of steps involving Cys-115 were attributed to a lowering of the  $pK_a$  value of the Cys-115 thiol. The Cys-115 thiolate was stabilized by interaction with the Lys-147 residue on the adjacent subunit. c) On the five-cysteine track, four hopping steps were observed at both +150 mV and -150 mV. The direction of the hops is indicated by the red arrows. Conditions: 2 M KCl, 50 mM HEPBS, pH 8.5, 20  $\mu$ M EDTA, 20  $\pm$  1 °C. d) Hopping rates involving Cys-113 and Cys-115 were accelerated under positive (113-to-115, 115-to-117) and negative potentials (117-to-115) on the E111Q-5C track relative to WT-5C track. Stepping rates for all steps were accelerated under both positive and negative potentials when the pH value of the buffer was increased to pH 9.5 (50 mM AMPSSO). Dwell-time analysis and rate constant estimations were performed by using the maximum interval likelihood algorithm of QuB software.<sup>[30]</sup> The average stepping rate along each track was calculated by taking the reciprocal of the average time per step (i.e., average time per step =  $\langle \tau_{total} \rangle / 8$ ).

5C)<sub>1</sub>WT<sub>6</sub> nanoreactor containing the engineered protein track E111Q-M113C-T115C-T117C-G119C-N121C was tested with a molecular hopper carrying an oligo-adenosine 40-mer cargo (Figure 4a).<sup>[9]</sup> The results were compared to the (WT-5C)<sub>1</sub>WT<sub>6</sub> nanoreactor containing the protein track M113C-T115C-T117C-G119C-N121C (Figure 4a). On the cysteine track, the hopper moved forward in the direction of Cys-113 to Cys-121 under +150 mV, and from Cys-121 to Cys-113 under -150 mV. Occasional backstepping could occur when the hopper moved in the reverse direction of that determined by the orientation of the DNA cargo. Both forward and backward steps occur by thiol-disulfide interchange, where the final foothold acts as the nucleophile, the initial foothold acts as the leaving group, and the hopper sulfur atom acts as an electrophile.

Oligonucleotide translocation in (E111Q-5C)<sub>1</sub>WT<sub>6</sub> was first investigated at pH 8.5. We observed faster 113-to-115 stepping under +150 mV ( $k_{113-115}^f = 0.12 \pm 0.02 \text{ s}^{-1}$  in (E111Q-5C)<sub>1</sub>WT<sub>6</sub> compared to  $0.044 \pm 0.009 \text{ s}^{-1}$  in (WT-5C)<sub>1</sub>WT<sub>6</sub>), while 115-to-113 stepping under -150 mV remained slow ( $k_{115-113}^b = 0.068 \pm 0.012 \text{ s}^{-1}$  in (E111Q-5C)<sub>1</sub>WT<sub>6</sub> compared to  $0.054 \pm 0.010 \text{ s}^{-1}$  in (WT-5C)<sub>1</sub>WT<sub>6</sub>; Figure 4d). We attribute the variations in rate enhancement to differences in DNA conformational lability under opposite potentials. When the DNA is bonded to Cys-113 and oriented towards the *trans* opening of the pore at +150 mV, about 4 phosphodiester groups lie within the electric field, which for the most part drops along the length of the  $\beta$  barrel. In contrast, an oligonucleotide bonded to Cys-115 and oriented towards the *cis* opening of the pore at -150 mV would only have about 1 phosphodiester group within the electric field.<sup>[9]</sup> Therefore, the stronger electrophoretic force acting on the DNA under +150 mV might align the disulfide bond linking the oligonucleotide to Cys-113 with Cys-115; by contrast, the oligonucleotide bonded to Cys-115 might have more degrees of freedom under -150 mV, which reduces occupation of the collinear conformation required for stepping to Cys-113.<sup>[29]</sup> Previous computational studies have shown that deviation of the collinear transition state from a S-S-S angle of 172° to 108° leads to an increase in activation energy of 167 kJ mol<sup>-1</sup>.<sup>[33]</sup> As predicted by PyMOL modeling, we also recorded an approximately 2-fold increase in the rate for the 115-to-117 step under +150 mV and for the 117-to-115 step under -150 mV in (E111Q-5C)<sub>1</sub>WT<sub>6</sub> (Figure 4d). The former is attributed to the Cys-115 thiolate being a better leaving group, and the latter to increased Cys-115 thiol deprotonation. All other steps occurred at similar rates by comparison to those in (WT-5C)<sub>1</sub>WT<sub>6</sub>.

We next studied oligonucleotide translocation at pH 9.5 in (E111Q-5C)<sub>1</sub>WT<sub>6</sub>. The rates of individual hopping steps at pH 9.5 increased by 5- to 8-fold under +150 mV and 2- to 3-fold under -150 mV by comparison with those measured at pH 8.5. Using the Henderson-Hasselbalch equation with the pK<sub>a</sub> values derived for individual cysteine residues, we predicted a 4- to 6-fold increase in stepping rates based on increased thiol deprotonation, which was in agreement with the observed rates under +150 mV. Smaller rate increases for steps under a potential of -150 mV could again be

attributed to the DNA having more degrees of freedom, thereby reducing occupation of the linear conformation required for thiol-disulfide interchange. While rate enhancement for the 115-to-113 step under -150 mV was low, the directionality of translocation was unaffected by 115-to-117 backstepping, with the rate of forward stepping ( $k_{115-113}^f = 0.12 \pm 0.01$ ) being an order of magnitude greater than that of backstepping ( $k_{115-117}^b = 0.012 \pm 0.005$ ) in (E111Q-5C)<sub>1</sub>WT<sub>6</sub>. Rates for other backsteps were not calculated owing to insufficient numbers of events (i.e., 14 backsteps out of 654 forward steps in (E111Q-5C)<sub>1</sub>WT<sub>6</sub> at pH 9.5, 7 of which were 115-to-117 backsteps under -150 mV).

For enzymeless sequencing, the biopolymer has to be moved from one end of the track to the other repeatedly for the measurement of current blockades. The mean total transit time ( $\langle \tau_{\text{total}} \rangle$ ) is defined as the time taken to complete a full cycle along a given track (i.e., from Cys-113 to Cys-121 under +150 mV followed by Cys-121 to Cys-113 under -150 mV). At pH 8.5, the E111Q mutation reduced  $\langle \tau_{\text{total}} \rangle$  from  $59 \pm 13 \text{ s}$  in (WT-5C)<sub>1</sub>WT<sub>6</sub> to  $40 \pm 9 \text{ s}$  in (E111Q-5C)<sub>1</sub>WT<sub>6</sub>. In particular, the E111Q mutation reduced the mean dwell time of the slowest step at pH 8.5 (i.e., 113-to-115 under +150 mV) from  $23 \pm 5 \text{ s}$  to  $8.4 \pm 1.7 \text{ s}$ , accounting for 70 % of the overall transit time reduction. If a uniform track with no hold-ups were desired, site-directed mutagenesis provides a means of targeting the slowest steps. A pH change from 8.5 to 9.5 further shortens  $\langle \tau_{\text{total}} \rangle$  to  $14 \pm 2 \text{ s}$  in (E111Q-5C)<sub>1</sub>WT<sub>6</sub>. Rate acceleration of the two slowest steps (i.e., 113-to-115 under +150 mV and 115-to-113 under -150 mV) accounted for 56 % of the overall transit time reduction. The average stepping rate was  $0.13 \pm 0.03 \text{ s}^{-1}$  in (WT-5C)<sub>1</sub>WT<sub>6</sub> at pH 8.5, and  $0.21 \pm 0.04 \text{ s}^{-1}$  and  $0.57 \pm 0.06 \text{ s}^{-1}$  in (E111Q-5C)<sub>1</sub>WT<sub>6</sub> at pH 8.5 and pH 9.5, respectively (Figure 4d).

## Conclusion

In conclusion, the mean total transit time for DNA cargo translocation was shortened from 59 to 14 s. Each step in (E111Q-5C)<sub>1</sub>WT<sub>6</sub> at pH 9.5 averaged  $0.57 \text{ s}^{-1}$ , a 4-fold increase in rate from the original system (i.e., (WT-5C)<sub>1</sub>WT<sub>6</sub> at pH 8.5). The rate-limiting step in the (WT-5C)<sub>1</sub>WT<sub>6</sub> nanoreactor was accelerated 21-fold ( $k_{113-115}^f = 0.044 \pm 0.009 \text{ s}^{-1}$  in (WT-5C)<sub>1</sub>WT<sub>6</sub> at pH 8.5 as compared to  $0.94 \pm 0.10 \text{ s}^{-1}$  in (E111Q-5C)<sub>1</sub>WT<sub>6</sub> at pH 9.5).

Our results from site-directed mutagenesis indicate that favorable changes to cysteine reactivity can be elicited by introducing the E111Q mutation. In this study, a 3.5- to 5.5-fold increase in reactivity was observed towards the small-molecule disulfide DTNB, while a 2- to 3-fold increase in reactivity was observed for selected steps on the cysteine track. All hopping steps within the (E111Q-5C)<sub>1</sub>WT<sub>6</sub> nanoreactor were sped up by increasing the pH value from 8.5 to 9.5. Without further track engineering, faster translocation could be achieved at a pH value above 9.5, where all cysteine footholds are fully deprotonated. We observed asymmetric rate enhancements under +150 mV and

–150 mV in the presence of the E111Q mutation and the pH increase. We speculate that the DNA cargo adopts different conformations under opposite potentials; the collinear transition state for thiol-disulfide interchange will be accessed less often when the DNA cargo has more degrees of freedom.<sup>[29]</sup>

The kinetics of mobile molecules might also be influenced by increasing the temperature. Previously, protein pores leukocidin, OmpG and various  $\alpha$ HL mutants were shown to be stable in single-channel recordings when the temperature was brought to more than 90 °C.<sup>[34]</sup> By this means, the binding of  $\beta$ -cyclodextrin to the  $\alpha$ HL mutant (M113N)<sub>7</sub> was shown to follow the linear form of the Van't Hoff equation. Using an activation energy of 50 kJ mol<sup>-1</sup>,<sup>[35]</sup> a 50 °C increase in temperature from 25 to 75 °C is predicted to produce an 18-fold rate acceleration for thiol-disulfide interchange. Assuming each step is accelerated 18-fold, the mean total transit time in (E111Q-5C)<sub>1</sub>WT<sub>6</sub> would be reduced to 0.78 s. Alternatively, selenocysteines might be incorporated into the protein track in place of cysteine. Selenolate (i.e., a deprotonated selenol) would act as the nucleophile and the leaving group during each hopping step. If selenium were incorporated into the biopolymer cargo, selenium could act as the electrophile for nucleophilic attack during each hopping step. Previous studies have suggested that 1) selenolates are approximately 1 order of magnitude more reactive than thiolates as nucleophiles; 2) selenium atoms are approximately 4 orders of magnitude more reactive than sulfur atoms as electrophiles.<sup>[36]</sup> Incorporation of an electrophilic selenium as the cargo handle could greatly reduce the mean total transit time to <1 ms, equivalent to >1000 complete cycles across the track per second.

Our system can be generalized for the translocation of biopolymers other than DNA, such as polypeptides. While the rates in our system do not so far compare to those of the helicase motors used in nanopore DNA/RNA sequencing devices, which make hundreds of steps per second, equivalent enzymes do not exist for other biopolymers, making sequencing by chemical stepping an attractive option. Furthermore, with optimization of conditions (i.e., increased pH value and temperature, the use of selenium, etc.), chemical stepping along cysteine tracks could be improved to provide rates comparable to those seen for sequencing devices for nucleic acids. We believe an ideal track for biopolymer translocation should have the following properties: 1) individual steps should be observable, and each foothold should be differentiated by the ionic current; 2) each track should support repeated translocation; 3) the track should be readily tunable by site-directed mutagenesis or chemical modification; 4) stepping rates should be similar throughout the track to avoid hold-ups during translocation. The present study presents significant advances in our endeavor to build an ideal track.

## Acknowledgements

This research was supported by the Wellcome Leap Delta Tissue Program, Oxford Nanopore Technologies, and a European Research Council Advanced Grant. Z.B. acknowledges the EPSRC Centre for Doctoral Training in Synthesis for Biology and Medicine for a studentship (EP/L015838/1), generously supported by AstraZeneca, Diamond Light Source, Defence Science and Technology Laboratory, Evotec, GlaxoSmithKline, Janssen, Novartis, Pfizer, Syngenta, Takeda, UCB and Vertex. Z.B. also acknowledges the China Scholarship Council-University of Oxford Scholarship for financial support. Z.B. thanks Dr T. K. Piskorz for helpful discussions on molecular dynamics simulations and H. Zhang for assistance on the linear regression analysis. Z.H.L. is grateful to the Centre for Doctoral Training in Synthesis for Biology and Medicine for a studentship, generously supported by GlaxoSmithKline, MSD, Syngenta, and Vertex. Z.H.L. also thanks the Oxford-Richards Scholarship for a studentship. Y.Q. was supported by a Glasstone Research Fellowship and a Fellowship by Examination, Magdalen College, Oxford, U.K. The molecular dynamics simulations in this work made use of time on JADE2 granted through the UK High-End Computing Consortium for Biomolecular Simulation, HECBioSim (<http://hecbiosim.ac.uk>), supported by the EPSRC (EP/T022205/1).

## Conflict of Interest

Hagan Bayley is the Founder of, a consultant for and a shareholder of Oxford Nanopore Technologies, a company engaged in the development of nanopore sensing and sequencing technologies.

## Data Availability Statement

The data that support the findings of this study are available from the corresponding authors upon reasonable request.

**Keywords:** Mobile Molecules · Nanopores · Protein Engineering · Single-Molecule Chemistry · Thiol-Disulfide Interchange

- [1] J. M. Berg, J. L. Tymoczko, G. J. Gatto, L. Stryer, *Biochemistry*, Macmillan International Higher Education, New York, **2019**, p. 2082.
- [2] H. Gu, J. Chao, S. J. Xiao, N.C. Seeman, *Nature* **2010**, *465*, 202–205.
- [3] K. Lund, A. J. Manzo, N. Dabby, N. Michelotti, A. Johnson-Buck, J. Nangreave, S. Taylor, R. Pei, M. N. Stojanovic, N. G. Walter, E. Winfree, H. Yan, *Nature* **2010**, *465*, 206–209.
- [4] S. J. Green, J. Bath, A. J. Turberfield, *Phys. Rev. Lett.* **2008**, *101*, 238101.
- [5] M. Von Delius, E. M. Geertsema, D. A. Leigh, *Nat. Chem.* **2010**, *2*, 96–101.

- [6] M. R. Wilson, J. Solà, A. Carlone, S. M. Goldup, N. Lebras-seur, D. A. Leigh, *Nature* **2016**, *534*, 235–240.
- [7] W. R. Browne, B. L. Feringa, *Nat. Nanotechnol.* **2006**, *1*, 25–35.
- [8] G. S. Pulcu, E. Mikhailova, L. S. Choi, H. Bayley, *Nat. Nanotechnol.* **2015**, *10*, 76–83.
- [9] Y. Qing, S. A. Ionescu, G. S. Pulcu, H. Bayley, *Science* **2018**, *361*, 908–912.
- [10] L. B. Poole, *Free Radical Biol. Med.* **2015**, *80*, 148–157.
- [11] S. H. Shin, T. Luchian, S. Cheley, O. Braha, H. Bayley, *Angew. Chem. Int. Ed.* **2002**, *41*, 3707–3709.
- [12] Y. Qing, H. Bayley, *J. Am. Chem. Soc.* **2021**, *143*, 18181–18187.
- [13] T. Luchian, S. H. Shin, H. Bayley, *Angew. Chem. Int. Ed.* **2003**, *42*, 3766–3771.
- [14] Y. Qing, G. S. Pulcu, N. A. W. Bell, H. Bayley, *Angew. Chem. Int. Ed.* **2018**, *57*, 1218–1221.
- [15] S. W. Englander, N. W. Downer, H. Teitelbaum, *Annu. Rev. Biochem.* **1972**, *41*, 903–924.
- [16] E. Liepinsh, G. Otting, *Magn. Reson. Med.* **1996**, *35*, 30–42.
- [17] G. Gao, E. F. DeRose, T. W. Kirby, R. E. London, *Biochemistry* **2006**, *45*, 1785–1794.
- [18] J. H. Tomlinson, S. Ullah, P. E. Hansen, M. P. Williamson, *J. Am. Chem. Soc.* **2009**, *131*, 4674–4684.
- [19] G. Maglia, A. J. Heron, D. Stoddart, D. Japrun, H. Bayley, *Methods in Enzymology*, Elsevier, Amsterdam, **2010**, pp. 591–623.
- [20] R. Singh, G. M. Whitesides, in *Sulphur-Containing Funct. Groups*, Wiley, Chichester, **1993**, pp. 633–658.
- [21] G. M. Whitesides, E. Jennifer, P. Szajewski, *J. Org. Chem.* **1977**, *42*, 332–338.
- [22] J. M. Wilson, R. J. Bayer, D. J. Hupe, *J. Am. Chem. Soc.* **1977**, *99*, 7922–7926.
- [23] P. Eyer, D. Podhradský, *Anal. Biochem.* **1986**, *153*, 57–66.
- [24] K. S. Jensen, J. T. Pedersen, J. R. Winther, K. Teilum, *Biochemistry* **2014**, *53*, 2533–2540.
- [25] E. V. Anslyn, D. A. Dougherty, *Modern Physical Organic Chemistry*, University Science Books, Sausalito, **2006**, p. 464.
- [26] R. M. LoPachin, T. Gavin, *Free Radical Res.* **2016**, *50*, 195–205.
- [27] G. Bulaj, T. Kortemme, D. P. Goldenberg, *Biochemistry* **1998**, *37*, 8965–8972.
- [28] M. B. Steffensen, D. Rotem, H. Bayley, *Nat. Chem.* **2014**, *6*, 603–607.
- [29] P. A. Fernandes, M. J. Ramos, *Chem. Eur. J.* **2004**, *10*, 257–266.
- [30] F. Qin, A. Auerbach, F. Sachs, *Biophys. J.* **1996**, *70*, 264–280.
- [31] P. F. Wang, M. J. McLeish, M. M. Kneen, G. Lee, G. L. Kenyon, *Biochemistry* **2001**, *40*, 11698–11705.
- [32] R. E. Hansen, H. Østergaard, J. R. Winther, *Biochemistry* **2005**, *44*, 5899–5906.
- [33] R. P. P. Neves, P. A. Fernandes, M. J. Ramos, *Proc. Natl. Acad. Sci. USA* **2017**, *114*, E4724–E4733.
- [34] X. F. Kang, L. Q. Gu, S. Cheley, H. Bayley, *Angew. Chem. Int. Ed.* **2005**, *44*, 1495–1499.
- [35] S. Chandrasekhar, B. S. Moorthy, R. Xie, E. M. Topp, *Pharm. Res.* **2016**, *33*, 1370–1382.
- [36] D. Steinmann, T. Nauser, W. H. Koppenol, *J. Org. Chem.* **2010**, *75*, 6696–6699.

Manuscript received: January 17, 2023

Accepted manuscript online: March 17, 2023

Version of record online: April 13, 2023