Single-molecule thiol-disulfide interchange



Yujia Qing

Brasenose College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

October 2018

Abstract

Thiol-disulfide interchange occurs widely in cellular processes including redox sensing and homeostasis, protein folding, cell signalling, and the regulation of apoptosis. Nature exploits this chemistry in a highly selective manner, which is challenging to recapitulate *in vitro*. In this thesis, we report site selective and regioselective thiol-disulfide interchange on macromolecular disulfide substrates elongated within a protein nanoreactor, where they react with cysteine thiolates presented at different locations along the length of a β strand. Numerous individual reaction events were detected by promoting the substrate turnover. For each substrate, we defined the most reactive cysteines on the β strand and which sulfur atom in the disulfide was attacked and found that the chemistry can be controlled with atomic precision.

We further applied our control over the selectivity of thiol-disulfide interchange to the development of a molecular machine. Intrigued by technological potential, scientists have long attempted to control molecular motion. We monitor the individual 0.7-nm steps of a single molecular hopper as it moves in an electric field along a track in a nanopore controlled by a chemical ratchet based on thiol-disulfide interchange. The hopper demonstrates characteristics desired in a moving molecule: defined start- and end-points, processivity, fuel autonomy, directional motion and external control. The hopper is readily functionalized to carry cargos. For example, DNA can be ratcheted along the track in either direction, a prerequisite for nanopore sequencing.

Publication list

1. Y. Qing, S. A. Ionescu, G. S. Pulcu, H. Bayley, Directional control of a processive molecular hopper. *Science*. **361**, 908–912 (2018).

2. Y. Qing, G. S. Pulcu, N. A. W. Bell, H. Bayley, Bioorthogonal cycloadditions with sub-millisecond intermediates. *Angew. Chem. Int. Ed.* **57**, 1218–1221 (2018). (Appendix I)

Acknowledgements

I would like to thank the following people.

Prof. Hagan Bayley, for his support and guidance over the years to shape me into a scientist; for telling me that ‘Progress is not linear’ when things were not working, which gladly turned out to be true; and for being an advocate of hard work and proper usage of the English language. This work would not be possible without his vision, pursuit of perfection in every aspect, and Saturday working hours.

The Bayley group, for the fond memories. In particular, Sandra Ionescu, for being a great friend to share a desk, a fridge shelf, snacks, lab necessities; for reading and editing my writings; and for staying inspirational. Nobina Mukherjee, for all her advice on scientific and non-scientific issues; for the fun holiday excursions taken together; and for being there at all times; Corey Kaminsky, Hiroko Tamagaki, Zonghua Bo, and Juan Liu for their infectious passion for science and life; and Michael Booth, for being a great colleague and a constructive critic, who also sides with me against lab mess, alcohol, and direct sunshine.

Prof. Mark Howarth, for the invention and generous supply of traptavidin for preliminary tests. Hiroko Tamagaki and Joongoo Lee, for their help with peptide design and synthesis; Lapatrada Taemaitree, for her help with LC-MS characterisation; Sejeong Lee, for her help with data analysis using QuB; Mira Liu for her considerable input on a project not presented in this thesis; Yiqun Geng, Qiong Zhang, and Yisi Lu, for motivating me with their stories behind the hard-earned achievements in their fields.

Lastly, my parents, for their unconditional support.

Abbreviations

αHL	α-Hemolysin
AFM	Atomic force microscopy
Cys	Cysteine
DIB	Droplet interface bilayer
DPhPC	1,2-Diphytanoly-sn-glycero-3-phosphocholine
dS	1′,2′-Dideoxyribose
DTNB	5,5′-Dithio(2-nitrobenzoic acid)
DTT	1,4-Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
GSH	Glutathione
GSSG	Oxidised glutathione
GUV	Giant unilamellar vesicle
HEPBS	N-(2-hydroxyethyl)piperazine-N′-(4-butanesulfonic acid)
Ires%	Residual current levels
KCl	Potassium chloride
MEE	Mercaptoethylether
PDI	Protein disulfide isomerase
PEG	Poly(ethylene glycol)
S_N2	Bimolecular nucleophilic substitution
TNB	5-Thio-2-nitrobenzoic acid

Contents

Abstract I

Publication list II

Acknowledgements III

Abbreviations IV

Contents V

Chapter 1 Introduction 1

1.1 Thiol-disulfide interchange in solution 1
1.2 Thiol-disulfide interchange in biology 6
1.3 Single-molecule studies of thiol-disulfide interchange 8
1.4 Scopes and objectives 16

Chapter 2 Regioselective thiol-disulfide interchange 17

2.1 Introduction 17
2.2 Turnover of thiol-disulfide interchanges 20
2.3 Controlled reaction between the nanoreactor thiolate and the substrate disulfide 21
2.4 Sulfur atom selection within the substrate disulfide 25
2.5 Kinetics of thiol-disulfide interchange 30
2.6 Site selective and regioselective tandem thiol-disulfide interchanges 32
2.7 Regioselectivity of thiol-disulfide interchange between DTT and αHL adducts 37
2.8 Discussions 40
2.9 Supplementary materials and methods 43

Chapter 3 Molecular Hopper 53

3.1 Introduction 53
3.2 Directional control of a processive molecular hopper 56
3.3 Hopping kinetics 62
3.4 Energetics of the directional movement 75
3.5 Releasing the hopper from a predefined foothold 77
3.6 Oligonucleotide sequencing 78
3.7 Supplementary materials and methods 83

Chapter 4 Conclusions and Perspectives 92

4.1 Conclusions 92
4.2 Perspectives and future directions 94
4.3 Outlooks 99

References 100

Chapter 1 Introduction

Thiols are ubiquitous in biology, and mediate important cellular functions (*1*, *2*). For example, the intracellular pool of thiols and disulfides maintain the redox homeostasis for antioxidant defence and redox regulation of cell signalling (*3*). The selective formation of structural disulfide bonds are necessary for the correct folding of proteins entering the secretory pathway (*4*). Cysteine-dependent catalysis is widely carried out by hydrolases, oxidoreductases, transferases, and isomerases (*5*). In contrast to the diversity of biological functions, these processes rely on simple thiol chemistries—thiol-disulfide interchange and thiol oxidation (*1*)—which often involve rapid and reversible formation of a covalent bond with the sulfur atom, and a high level of selectivity (*6*). Here, we will review the efforts put into understanding thiol-disulfide interchange, a unique thiol chemistry of pivotal role in biology.

1.1 Thiol-disulfide interchange in solution

Thiols (RSH) are organosulfur compounds containing a carbon-bonded sulfhydryl group (-SH), which exhibit reactivity towards electrophiles and oxidants, and have high affinities for metals (*1*–*3*). Deprotonation of a thiol generates a thiolate anion (RS^-) with stronger nucleophilicity, and the degree of thiol ionisation at a given pH can be inferred from the associated pK_a value (*7*). The pK_a of a thiol varies depending on the molecular environment, which is ~8.5 for small-molecule thiols such as glutathione (GSH) and unperturbed cysteines in proteins (*8*).

Thiols can undergo oxidation to form a disulfide (RSSR) (Eq. 1).

$$2\ RSH \leftrightharpoons RSSR + 2H^+ + 2e^- \qquad \text{Eq. 1}$$

The disulfide exhibits an S-S bond energy of 251 kJ mol^{-1} and a bond length of 2.05 Å, which is weaker than a C-C bond (348 kJ mol^{-1}; 1.54 Å) (*7*). A disulfide prefers

a C-S-S-C dihedral angle of ~90° (Fig. 1.1A) with a rotational barrier around the S−S axis of ~30 kJ mol^{-1} to the cis conformer, and ~20 kJ mol^{-1} to the trans conformer (*9*).

A reversible reaction between a thiol (R_1SH) and a disulfide (R_2SSR_3)—known as a thiol-disulfide interchange—generates a new disulfide (R_1SSR_2) and releases a thiol from the original disulfide (R_3SH) (Eq. 2-4). This S_N2 reaction involves backside attack of the nucleophilic thiolate along the S-S axis of the disulfide (*10–12*). Upon interchange, the negative charge delocalises over all three sulfur atoms in a trisulfide-like transition state (S-S-S) (Fig. 1.2, reviewed later in this section) (*11–14*).

The associated rate of reaction is second-order overall—first-order in both thiolate and disulfide (*7*, *15*, *16*). The observed rate constant (k_{obs}) is therefore dependent on both the thiol pK_a and the solution pH (Eq. 5) (*7*).

$$R_1SH \leftrightharpoons R_1S^- + H^+ \quad \text{Eq. 2}$$

$$R_1S^- + R_2SSR_3 \leftrightharpoons R_1SSR_2 + R_3S^- \quad \text{Eq. 3}$$

$$H^+ + R_3S^- \leftrightharpoons HSR_3 \quad \text{Eq. 4}$$

$$k_{obs} = k_{RS^-} / (1+10^{pKa-pH}) \quad \text{Eq. 5}$$

where k_{RS^-} is the pH-independent rate of the thiolate-disulfide interchange (Eq. 3).

A series of empirical studies using small-molecule thiols established a linear relationship between the log of the rate of the thiolate-disulfide interchange ($\log(k_{RS^-})$) and the pK_a of the nucleophilic thiols (Fig. 1.1B). This fits into a Brønsted correlation for a base-catalysed reaction (Eq. 6), where β_{nuc} (slope) is the Brønsted coefficient for the nucleophilic thiols (*11*, *17*, *18*). The thiol pK_a value which leads to a maximum value of k_{obs} at a given pH can be in turn predicted following Eq. 7 (*17*).

$$\log(k_{RS^-}) = (\beta_{nuc} \times pK_a) + C \quad \text{Eq. 6}$$

$$pK_a = pH - \log [(1- \beta_{nuc}) / \beta_{nuc}] \quad \text{Eq. 7}$$

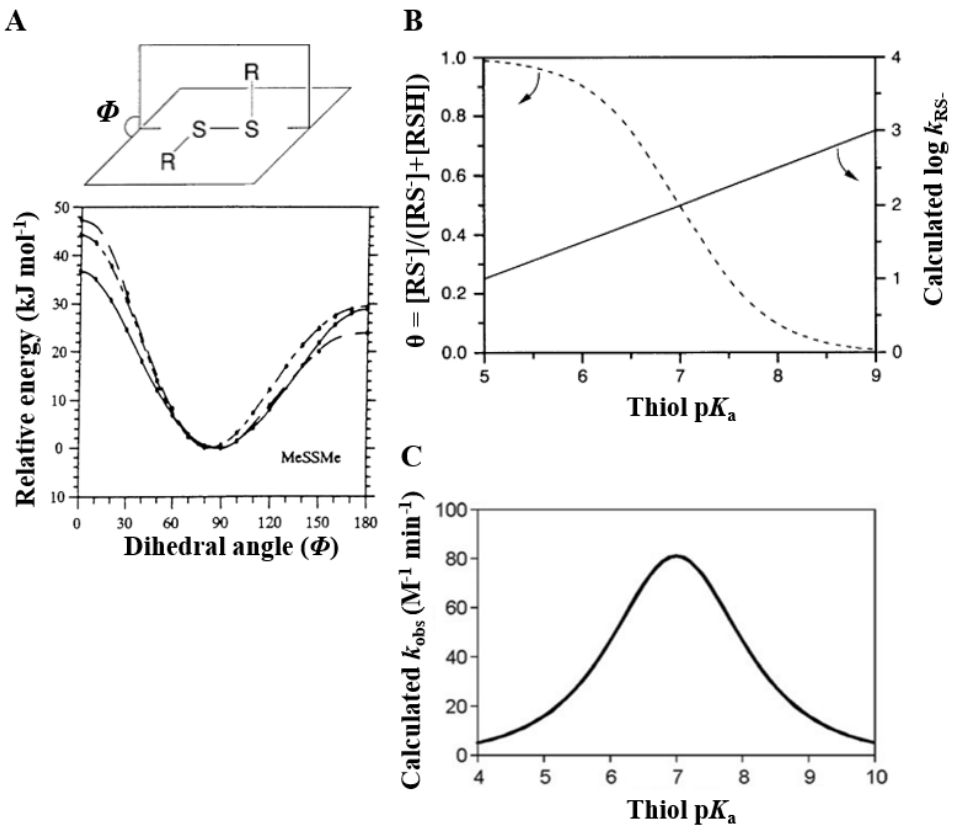


Fig. 1.1. Disulfide properties and pK_a dependence of kinetics of thiol-disulfide interchange. (**A**) Relative energy of MeSSMe as a function of the C-S-S-C dihedral angle (Φ). The most stable conformer has Φ = 90°. The cis conformer has Φ = 0°, while the trans conformer has Φ = 180°**.** The three plots were derived by three separate computational methods (*19*). (**B**) The log of k_{RS^-} was plotted as a function of thiol pK_a (solid line). The values of k_{RS^-} were calculated using the experimentally established Brønsted relationship for a thiol-disulfide interchange with β_{nuc} of 0.5, and pK_a values of 8.5 for both the central and leaving thiols. The degree of dissociation at pH 7 (θ) was plotted as a function of thiol pK_a (dashed line) for comparison (*7*). (**C**) Calculated k_{obs} for thiol-GSSG interchange as a function of thiol pK_a at pH 7 based on the Brønsted relationship established experimentally (β_{nuc} is 0.5). The k_{obs} peaked when thiol pK_a equalled pH, indicating that k_{obs} is dependent on both the pK_a and the nucleophilicity of the thiol (*6*).

In general, β_{nuc} for the small-molecule thiols that were tested ranged between 0.4 and 0.6 (*11*, *17*, *18*). Particularly, the thiol-disulfide interchange between thiolates and oxidised glutathione (GSSG) (*18*) or Ellman's disulfide (5,5′-dithiobis-(2-nitrobenzoic) acid, DTNB) (*11*, *17*) have β_{nuc} of 0.5, which suggests a maximum k_{obs} when a thiol has pK_a equal to pH (Fig. 1.1C). A simple explanation is that a thiol with

pK_a<<pH stays mostly in the deprotonated form but exhibits less nucleophilicity as a thiolate; a thiol with pK_a>>pH is less prone to ionisation albeit is a good nucleophile as a thiolate. The Brønsted analysis therefore provides a guideline for balancing the intrinsic nucleophilicity and the extent of ionisation for a given thiol to favour fast thiol-disulfide interchange.

Using either a thiolate or a disulfide presented by a polypeptide or a protein, a log k_{RS^-} versus pK_a Brønsted correlation was obtained in most cases, with β_{nuc} values similar to those seen with small-molecule compounds (*20*, *21*). However, it was found that the rate of thiol-disulfide interchange was also under the substantial influence of electrostatic effects associated with macromolecular hosts. For example, the β_{nuc} was ~0 for the reactions between GSSG and polypeptides with varying charges (*21*). In this case, favourable electrostatic interactions increased the probability of forming an initial encounter complex between the negatively charged GSSG and model peptides with net positive charge (and low thiol pK_a values due to the thiolate stabilisation by nearby positive residues) (*21*). This generated a negative correlation (i.e. low pK_a values correlate with fast rates, hence the slope of the Brønsted plot is negative) and offset the positive trend established with neutral reactants.

Similarly, a Brønsted correlation can be established between the k_{RS^-} and the pK_a values for the central (c) and leaving group (lg) thiols. The corresponding Brønsted coefficients, β_c and β_{lg}, were less well characterised and roughly estimated to be between -0.3 and -0.4 for β_c and -0.5 for β_{lg} based on a limited data set (*7*, *22*). Take 1,4-dithiothreitol (DTT) reduction of a symmetrical disulfide, for example: DTNB was reduced at k_{obs} = 2.5×10^3 $M^{-1}s^{-1}$ (pH 7.0, 30°) (*17*) whereas GSSG was reduced at k_{obs} = 0.59 $M^{-1}s^{-1}$ (pH 7.0, 30°) (*23*). The constituent thiol of DTNB (5-thio-2-nitrobenzoic acid, TNB) has a pK_a of 4.5 (*24*) and that of glutathione (GSH) is 8.7 (*23*).

Computational simulations of thiol-disulfide interchange in solution using model compounds revealed a transition state where the three participating sulfur atoms adopt a collinear arrangement during the S_N2 backside attack (*25–27*) (Fig. 1.2). In the transition state, both S-S attain intermediate lengths (~2.5 Å) between that of a single bond and a nonbonding contact (*25–27*) and charge is distributed among three atoms (Fig. 1.2A). A transition state of similar structure was also observed during in silico modelling of an enzymatic reduction of GSSG by a protein disulfide isomerase (PDI) (Fig. 1.2B) (*28*). This high degree of linear directionality intrinsic to the S_N2 thiol-disulfide interchange was later evidenced by single-molecule force spectroscopy studies (*29–31*) (reviewed in section 1.3.1).

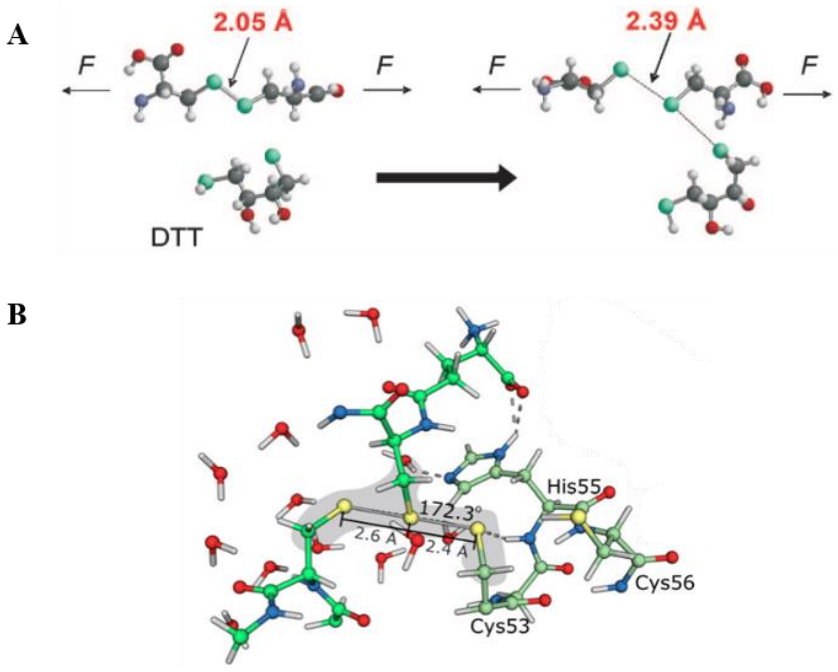


Fig. 1.2. Transition state during thiol-disulfide interchange. (**A**) Schematic representation of a S_N2 reaction between DTT and a protein disulfide under force. The transition state involves three sulfur atoms (green) aligned collinearly. An elongation of the disulfide bond by 0.34 Å was derived from the force-dependency of the disulfide reduction rate (*29*). (**B**) Computational model of a linear transition state for the first step of PDI reduction of GSSG. PDI active site residues are labelled and shown in pale green, GSSG is shown in lime green, and sulfurs are in yellow (*28*).

1.2 Thiol-disulfide interchange in biology

Thiol/disulfide chemistry is heavily employed by biology, with well-recognised roles in the maintenance of redox equilibria, the folding and stabilisation of proteins, enzymatic catalysis, and the regulation of protein activities (*32–35*). Fass and Thorpe have thoroughly reviewed numerous biologically relevant chemistries involving enzymes and small molecules (e.g. GSH, quinones, flavins), among which thiol-disulfide interchange remains the most prevalent (*1*). This thesis will focus on one aspect of thiol-disulfide interchange, which is well demonstrated in biology but less discussed in terms of chemistry: regioselectivity.

Regioselectivity, as broadly defined, presents two aspects with respect to thiol-disulfide interchange: the differentiation between two sulfurs within a single disulfide by an attacking thiolate, and the differentiation between multiple thiolates or disulfides at different loci by a reactive partner. Here, we will refer to the former as 'regioselectivity', in its narrow sense and the latter as 'site-selectivity' (*36*, *37*).

The structures of many redox proteins and enzymes have been solved at the stage where they have formed a specific mixed-disulfide linkage with a substrate (Fig. 1.3A) (*38–42*) or have switched to a configuration poised for sequential disulfide shuffling in a defined order (*43*, *44*). Given the presence of multiple cysteines in those proteins, the observations strongly suggested that thiol-disulfide interchange was implemented with high regioselectivity and site-selectivity at the captured steps (Fig. 1.3B). It has been concluded that dominating thiol-disulfide interchange pathway is largely directed by the conformations adopted by protein-protein interactions between the proteins involved (*1*). where steric influence prevails. This indicates that biology has evolved to exploit the selectivity of the chemistry, such that only access to the 'correct'

linear transition state occurs, and the geometric requirements of in-line attack are not fulfilled for other possible pathways (discussed in section 1.1).

As alluded to earlier (section 1.1), the reactivity of the involved thiol and disulfide determines the regioselectivity of interchange in solution. For example, the interchange between a thiolate and an asymmetrical disulfide (R_1SSR_2, pK_a: R_1SH > R_2SH) prefers the regioselective release of the more acidic thiol (R_2SH) (*22*). We speculate that when multiple sites (either thiols or disulfides) are available for thiol-disulfide interchange within a small molecule, their reactivity will influence the probability of reaction at each site, hence the site-selectivity. However, such intrinsic reactivity of thiols and disulfides is often missed from the discussions on thiol-disulfide interchange in biology even though its influence on reactivity can be on par with that of geometric constraints in the context of a protein.

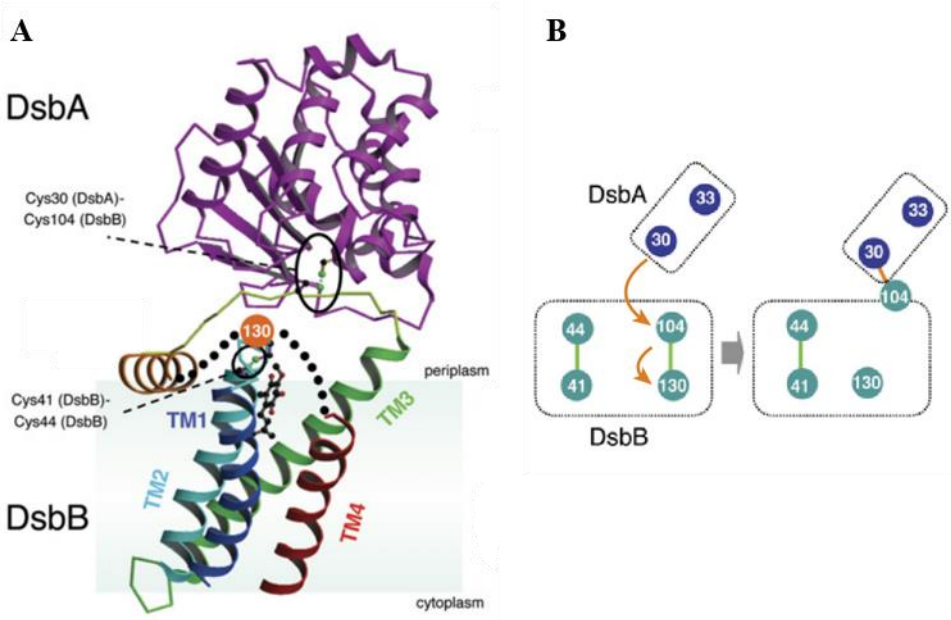


Fig. 1.3. A mixed-disulfide between Disulfide bond formation protein A (DsbA) and Disulfide bond formation protein B (DsbB) formed after a regioselective and site-selective thiol-disulfide interchange. (**A**) Ribbon representation of the DsbB-DsbA complex viewed parallel to the membrane plane. The Cys-30 (DsbA)-Cys-104 (DsbB) and Cys-41 (DsbB)-Cys-44 (DsbB) disulfide bonds are circled. (**B**) Schematic representation of the formation of DsbA-DsbB complex, where Cys-30 (DsbA) attacks Cys-104 (DsbB) regioselectively and site-selectively. (Reproduced from (*38*))

1.3 Single-molecule studies of thiol-disulfide interchange

With the advent of single-molecule techniques such as single-molecule force-spectroscopy and single-channel recording, atomistic details of the thiol-disulfide interchange have been uncovered in the context of both small molecules and macromolecules. At the single-molecule level, dynamic molecular behaviour is revealed, which is obscured in conventional ensemble studies by bulk averaging.

1.3.1 Single-molecule force-spectroscopy

In biology, thiol-disulfide interchange predominantly takes place on proteins, which provide complex reaction landscapes that differ from the conditions with small molecules in solution. Reactions between thiols and disulfides are subjected to the influence of not only the external conditions, such as the redox potential and mechanical stress, but also the local environment within the hosting macromolecules.

To explore thiol-disulfide interchange in the context of a protein at single-reaction resolution, Fernandez and Garcia-Manyes pioneered the use of single-molecule force-spectroscopy together with engineered disulfide bonds within a titin immunoglobulin domain (Fig. 1.4) (*45*, *46*). In a typical experiment, a polyprotein construct (e.g. an oligomer containing 8 repeats of an engineered titin domain (*47*, *48*)) was held between an atomic force microscopy (AFM) cantilever tip and a Au substrate attached to a piezoelectric actuator (Fig. 1.4A) (*45*, *46*). In this way, a force was applied with defined directionality and magnitude to unfold the protein up to the disulfide bond, which was then exposed to external reducing agents (*29*, *31*, *49*). Disulfide reduction (or reformation) was precisely correlated with changes in protein contour length during unfolding (or refolding) under an applied force (Fig. 1.4B) (*50*). Consequently, atomistic details of the thiol/disulfide chemistry were revealed from real-time display of the reaction dynamics at single-bond resolution (*29–31*, *51*). This approach has allowed

the dissection of catalytic mechanisms underlying redox proteins (*49*, *52*) and oxidoreductases (*48*, *53*). Non-enzymatic disulfide rearrangement on a titin domain mutant was also demonstrated (*47*, *54*), which resembled the PDI-triggered disulfide rearrangement on a substrate and exhibited surprising regioselectivity and site-selectivity (reviewed in details in **Chapter 2**).

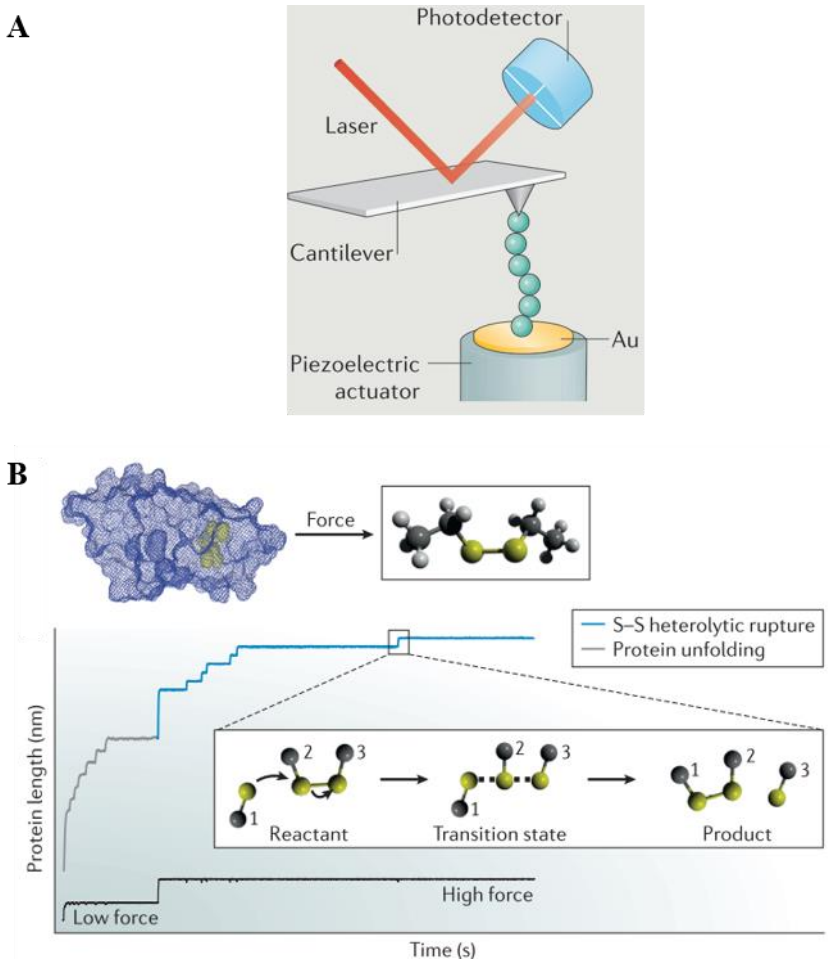


Fig. 1.4. Single-molecule force spectroscopy studies of thiol-disulfide interchange. (**A**) In the AFM setup, the target protein is held between two tethering surfaces, the distance between them can be manipulated to apply mechanical stress (~10–2,000 pN). Subsequent conformational changes are recorded as a change in the end-to-end length of the molecule (*45*). (**B**) In a single-molecule force-clamp AFM experiment, the disulfide bond engineered into the core of a titin domain was exposed during protein unfolding (grey steps) and subsequently reduced by a reducing reagent in solution (blue steps). Steps were assigned based on the step change in the protein length which corresponds to the extension in polypeptide chain after each event (*45*).

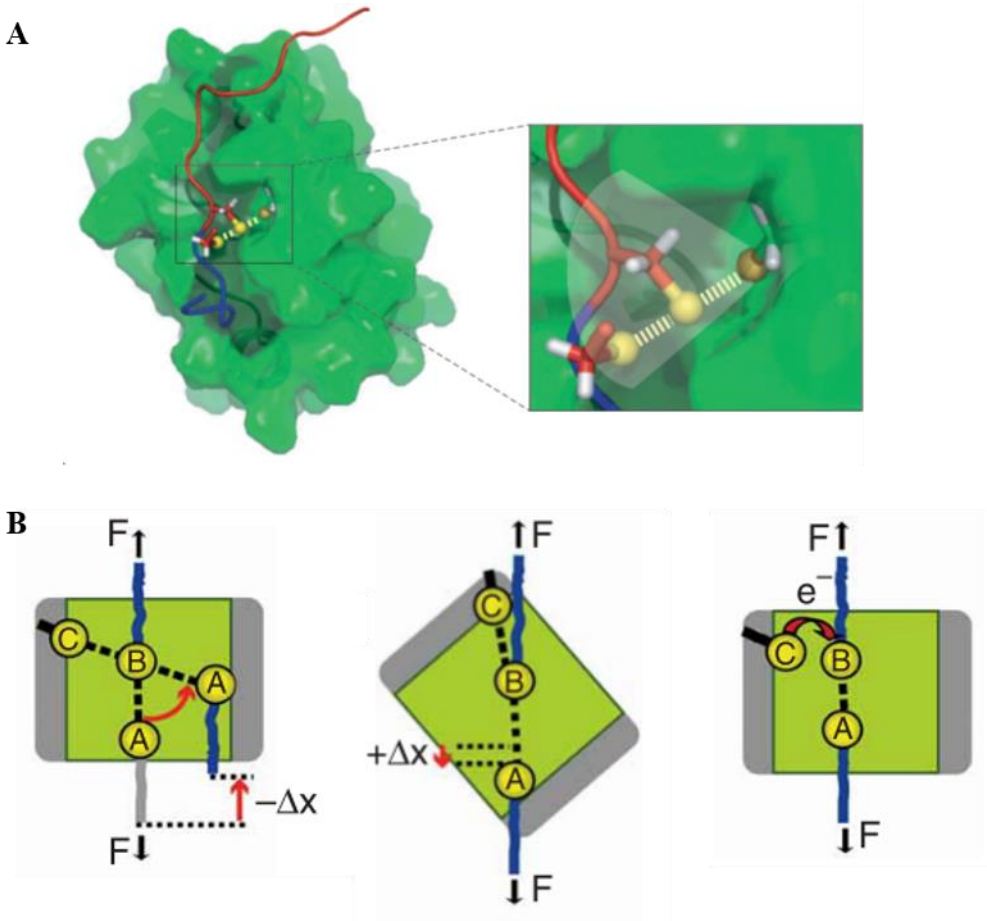


Fig. 1.5. Single-molecule spectroscopy studies of thiol-disulfide interchange. (**A**) Disulfide reduction by thioredoxin was probed under force, where conformational re-orientation of the disulfide was required in order to reach the linear transition state (first mechanism shown in (**B**)) (*49*). (**B**) Further studies of the thioredoxin-mediated disulfide reduction revealed three types of force-dependencies of the kinetics. They were attributed to three possible mechanisms of disulfide reduction under force: the disulfide re-orientation prior to S_N2 (left); the direct S_N2 (middle); the single-electron transfer mechanism via tunnelling (right) (*52*).

Specifically, during the mechanical unfolding of disulfide-containing titin domains in the presence of small-molecule thiols, the reduction of a single disulfide bond was found to follow an single exponential time-course distribution (*31*). This supported the notion that thiol-disulfide interchange involving small-molecule thiols undergoes a single transition state, which was assigned as the well-defined S-S-S Walden complex (*29*, *31*). In contrast, the time distribution of titin unfolding events

deviated significantly from a single exponential, which suggested heterogeneous transition states during unfolding (*31*).

On the other hand, the rate of disulfide reduction by small-molecule thiols was found to be exponentially dependent on the applied force, increasing 10-fold over a 300 pN range (*29*). The obtained correlation between the disulfide reduction rate and the applied force predicted a S-S bond lengthening of ~0.3-0.4 Å to reach transition state, which agreed well with ab initio calculations (*30*).

Conversely, complex disulfide reduction kinetics were observed with redox protein (e.g. thioredoxin) under force, which suggested multiple reaction trajectories (Fig. 1.5A) (*49*, *52*). Three disulfide reduction mechanisms were proposed, which involve a disulfide re-orientation step inside a protein binding groove prior to the S_N2 attack, or a direct S_N2 reaction when the substrate sits in a shallow binding groove and the disulfide is aligned with the cysteine, or a single-electron transfer mechanism that does not require the collinear arrangement of sulfur atoms (*49*, *52*) (Fig. 1.5B).

1.3.1.1 Nanopore approach to single-molecule thiol-disulfide interchange

The Bayley group has developed a means to monitor chemical reactions in aqueous solution at the single-molecule level within a protein nanoreactor (Fig. 1.6) (*55*). The nanoreactor is constructed from a heptameric pore-forming toxin, α-hemolysin (αHL) (Fig. 1.6), which inserts into a synthetic lipid bilayer and allows steady ionic current flow through its β barrel under an applied potential (*56*). To create a single reactive site, one of the seven subunits of the αHL is engineered with a reactive group facing the lumen of the β barrel. Conveniently, thiol chemistry can be easily studied with an inward-facing cysteine residue introduced into a single monomer by mutagenesis (*57–60*). To study other chemistries, the reactant has been tethered to the interior cysteine through a Michael-addition (*61*, *62*) or an alkylation with an α-

halocarbonyl electrophile (i.e. iodoacetamide) (*63–65*). Alternatively, native chemical ligation between a C-terminal thioester and an N-terminal cysteine has been used to incorporate unnatural amino acids into αHL for nanoreactor construction (*66*, *67*). Chemistry occurring inside the nanoreactor perturbs the ionic current passing through it, thereby allowing tag-free real-time detection of intermediate species formed during a single reaction and determination of the associated lifetimes. The temporal resolution of the approach can be in the millisecond range (i.e. rates of reaction >100 s^{-1} (*63*)), making it suitable for studying thiol-disulfide interchange. Indeed, the thiol-disulfide interchange between a nanoreactor cysteine, Ellman's reagent, and DTT has been characterised at the single-molecule level using the nanopore approach (*58*, *59*), as reviewed in the following section.

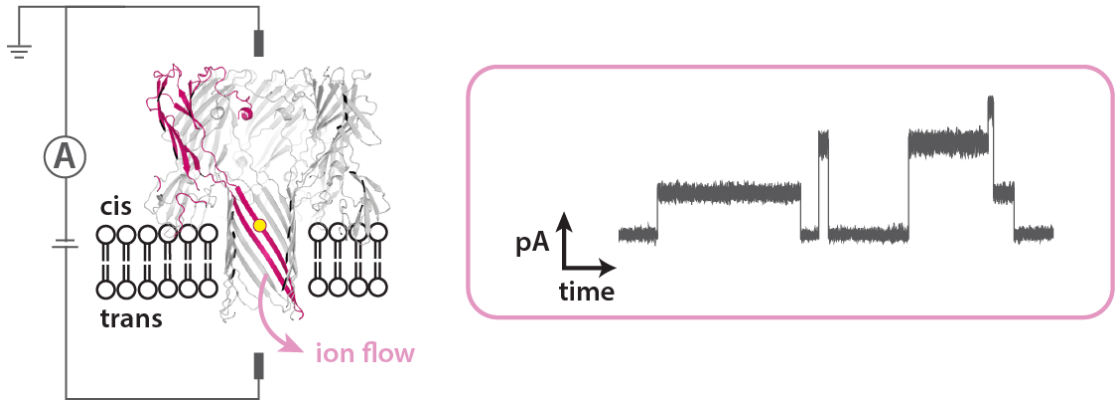


Fig. 1.6. Single-molecule chemistry is studied using single-channel recording. Under an applied potential, a single α-hemolysin (αHL) inserts into the synthetic lipid bilayer to allow transmembrane ion flow. A positive current is defined as cations flowing through the pore from trans to cis (or anions from cis to trans). A single reactive site (yellow circle) is engineered into one of the seven subunits of the nanopore (highlighted in pink). Chemical reactions at the reactive site perturb the ionic current, which is recorded in real-time for single-molecule chemistry studies (pink box).

1.3.1.2 Nanopore approach to turn-over in Ellman's chemistry

In solution, the thiol-disulfide interchange between DTT and Ellman's reagent (DTNB) is rapid and irreversible (*17*). To probe this chemistry at the single-molecule

level, the Bayley group utilised a synthetic lipid bilayer—in which a single-cysteine nanoreactor was embedded—as a physical barrier to prevent reactions between DTT and DTNB outside the nanoreactor (*58*). Spatial separation of DTT and DTNB enabled turnover of the thiol-disulfide interchange between them at the single cysteine site within the barrel (Fig. 1.7A). Typically, a cycle started with the cysteine sulfhydryl group attacking a DTNB molecule diffusing from the trans compartment, which resulted in a mixed-disulfide between TNB and nanoreactor cysteine (step 1, Fig. 1.7B and 7C). An incoming DTT from the cis compartment then attacked the disulfide sulfur originating from the nanoreactor cysteine, generating a transient αHL-DTT adduct (step 2, Fig. 1.7B and 7C)—previously undetected by single-molecule force spectroscopy (*29*). Spontaneous breakdown of this adduct was triggered by the self-cyclisation of DTT to form a six-membered ring (step 3, Fig. 1.7B and 7C).

The rate of reaction for each step was determined at different pH values as well as different DTT and DTNB concentrations (*58*). For example, the rate constant for step 3 (k_3) (Fig. 1.7B and 7C) was independent of the concentrations of DTT or DTNB but showed a clear dependence on pH. This was expected for an intramolecular thiol-disulfide interchange, in which a thiolate anion acts as the nucleophile. Based on single-channel recording data, a pK_a value of 9.5 ± 0.2 was derived for the free thiol within the αHL-DTT adduct and a self-cyclisation rate (k_3) of 250 ± 80 s^{-1} was determined for DTT in the deprotonated form. The kinetic profile of step 3 had never been reported before, presumably due to the lack of techniques to characterise short-lived intermediates in an ensemble experiment, especially when its breakdown is not the rate-limiting step (*55*).

It is worth noting that, the regioselectivity of step 2—where the mixed-disulfide is cleaved by the DTT thiolate—was biased completely towards the release of TNB,

and the cysteine sulfur became bonded to the DTT sulfur in the newly formed mixed-disulfide. The complete regioselectivity presumably arose from the TNB being an excellent leaving group, in keeping with previous ensemble observations (*22*). This work showcased the power of the single-channel recording technique to determine regioselectivity of an individual chemical reaction between two atoms within a single bond.

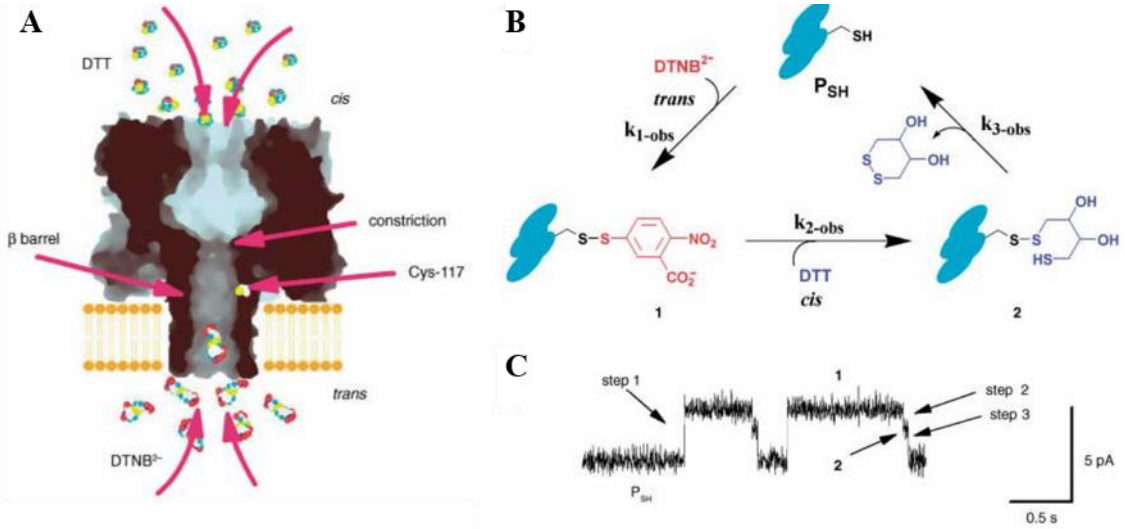


Fig. 1.7. Nanopore approach to study thiol-disulfide interchange at the single-molecule level. (**A**) A αHL nanoreactor containing a single cysteine at position 117 was inserted into the bilayer. The side presented by the mushroom cap of the nanoreactor is termed cis and the side presented by the opening of the β barrel is the trans side. DTT (cis) and DTNB (trans) were spatially separated (*58*). (**B**) Cycles of three consecutive thiol-disulfide interchanges were detected at the single-molecule level with the corresponding current signature shown in (**C**) (*58*). Conditions: 2 M KCl, 30 mM 3-(4-morpholinyl)-1-propanesulfonic acid (MOPS), 100 µM EDTA, pH 8.5, -50 mV (trans).

1.3.1.3 Stepwise polymer growth based on Ellman's chemistry

Using the same principles to turn over thiol-disulfide interchange described above (*58*), a system consisting of spatially separated DTNB and a dithiol reagent—(mercaptoethyl)-ether (MEE)—was constructed (*59*) (Fig. 1.8A). Unlike DTT, MEE self-cyclised at a significantly slower rate due to the unfavourable formation of the corresponding seven-membered ring. This allowed the free thiol within the αHL-MEE

adduct to further couple with a second DTNB (Fig. 1.8B). The resultant terminal activated disulfide subsequently reacted with another MEE molecule to form a MEE dimer (Fig. 1.8B). In this way, a polymer chain grew in steps and the associated extension rate for individual steps was determined (Fig. 1.8B) (*59*). Addition of up to 10 MEE molecules were detected, at which point the polymer fully blocked the ion flow through the nanoreactor (*59*). Overall, the average rate constant for the reactions with DTNB was 1.6×10^4 M^{-1} s^{-1} whereas the average rate for MEE addition was 2.7×10^3 M^{-1} s^{-1}. Complete regioselectivity was preserved for attack on the mixed-disulfide between TNB and the nanoreactor cysteine.

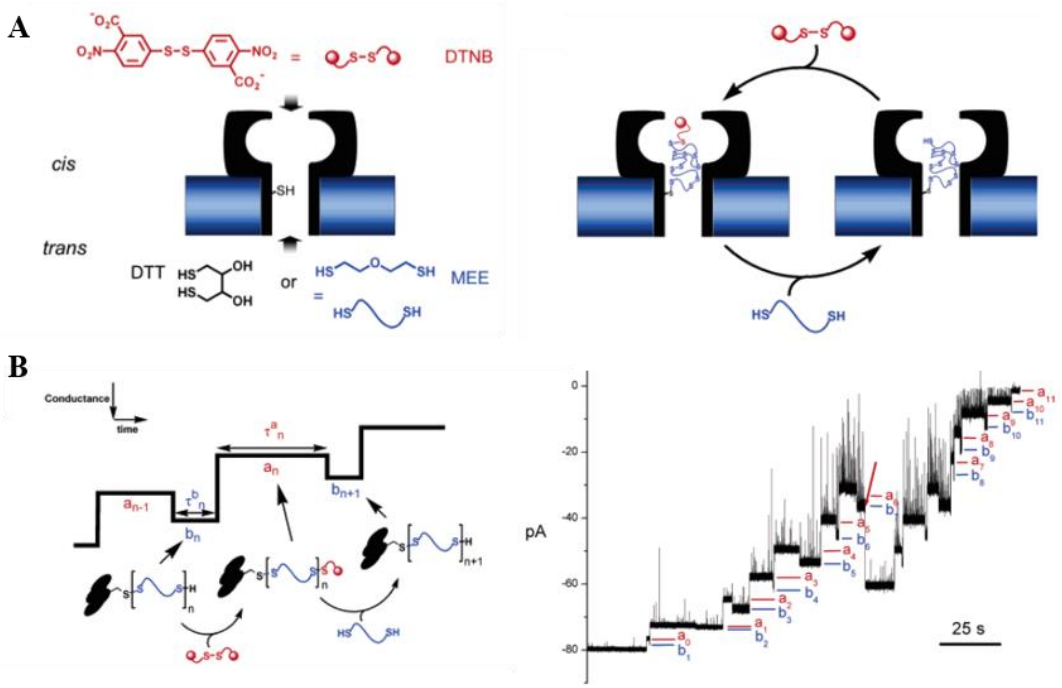


Fig. 1.8. Nanopore approach to study dynamic polymer growth based on thiol-disulfide interchange. (**A**) A single-cysteine αHL nanoreactor was used to probe thiol-disulfide interchange between spatially separated reactants. DTNB and MEE were added to cis and trans, respectively, and reacted to form a polymer attached at the cysteine residue (*59*). (**B**) The left scheme shows a schematic of the lifetimes (τ) and conductance levels of the intermediates—one with a terminal activated disulfide formed after DTNB coupling (red, a_n) and another containing a free thiol formed after MEE reduction (blue, b_n). A representative single-channel recording trace on the right shows the stepwise growth of the polymer chain (*59*). Conditions: 2 M KCl, 30 mM MOPS, 100 µM EDTA, pH 8.5, -50 mV (trans).

1.4 Scopes and objectives

Here we have reviewed the intensive efforts put into understanding thiol-disulfide interchange, which is a seemingly simple but versatile chemistry with a pivotal role in biology. Two single-molecule techniques have been applied to investigate this chemistry, which have revealed a wealth of details regarding its kinetics, transition state geometry, as well as its regioselectivity and site selectivity. Based on this knowledge, we envision that the chemistry of thiol-disulfide interchange is tunable and the ability to control it will open up a multitude of possible applications, including chemical signalling and molecular motion control. This thesis is centred around achieving these goals.

Chapter 2 presents a strategy to control the regioselectivity and site selectivity of thiol-disulfide interchange by using spatial alignment. The disulfide is placed within a macromolecule to realise geometric constraints, an approach inspired by the handling of macromolecules during single-molecule force spectroscopy. The thiolate is engineered within an αHL nanoreactor to permit the use of single-channel recording for the real-time detection of atom and site selection during individual interchange reactions. Future applications in chemical signalling within a synthetic tissue are also proposed.

Chapter 3 presents a processive molecular hopper ratcheted by thiol-disulfide interchange on a cysteine track within the nanopore. The directional control of the molecular motion and the ability to load and release the hopper at precisely determined sites rely on regioselective and site-selective control over the chemistry. The hopper carries a DNA cargo, which is moved in 0.7 nm steps by the hopping motion, which may provide an alternative ratcheting machinery for nanopore sequencing.

Chapter 4 summaries the findings and proposed future directions.

Chapter 2 Regioselective thiol-disulfide interchange

2.1 Introduction

The control of selectivity has been a longstanding challenge for synthetic chemists. Two important aspects of this endeavour are site-selectivity—the ability to differentiate between two (or more) similarly reactive positions within a molecule, and regioselectivity—the ability to distinguish between two (or more) sites within a given functional group (*36*, *37*). The realisation of selectivity often requires exploitation or subversion of the inherent steric and electronic properties of a molecule or functional group (*36*). Common strategies include the installation of a reagent-directing group on the substrate, which can be later removed (*68*), or the recruitment of a selective catalyst (organic or transition-metal-based), which interacts with the substrate to favour a particular reaction pathway. Alternatively, the preorganization of reactants, in micelles (*69*) or within a supramolecular cage (*70*), has been used to alter innate regioselectivity. Recently, site selectivity was demonstrated as a result of the spacing of reactive sites within a metal-organic framework (*71*). In the context of enzymology, selectivity has been exquisitely demonstrated (*72*). For example, a switch in the regioselectivity of nitration of L-tryptophan by a cytochrome P450 has been attributed to a change in the orientation of the substrate, which depended upon a single amino acid substitution (*73*). Despite these remarkable achievements, the intricate molecular interactions underlying many site selective or regioselective chemistries were only elucidated after their serendipitous discovery and in some cases are still not fully understood (*36*).

Thiol-disulfide chemistry is prevalent in nature where it underpins critical biological activities involving enzymes such as thioredoxins and protein disulfide isomerases. In the present work, we use a single-molecule approach to examine thiol-disulfide interchange. Previously, single-molecule force-clamp spectroscopy was used to study thiol-disulfide chemistry under mechanical stress by monitoring the unfolding of a titin domain mutant containing two disulfides (*47*). Unfolding was initiated by L-cysteine, which reduced the more exposed disulfide. Strikingly, under an applied force, the cysteine side chain liberated from the first disulfide reacted regioselectively with the remaining disulfide, preferring one sulfur atom ~4 times over the other. Subsequent computational simulations attributed the regioselectivity mainly to the polypeptide conformation after partial unfolding (*74*), which would be impossible to control in bulk solution. Until now, an artificial system in which a nucleophile is programmed to react with one of two chemically equivalent sulfur atoms in a disulfide has not been reported.

The system we report here comprises a nanoreactor—an α-hemolysin (αHL) cysteine mutant (Fig. 2.1A)—and a linear macromolecular substrate with a disulfide bond in its backbone (Fig. 2.1B). The substrate consisted of biotin—a traptavidin-binding motif—at the terminus of a peptide segment, which was conjugated to a DNA oligonucleotide by strain-promoted azide-alkyne cycloaddition. The disulfide was installed between the peptide and the oligonucleotide with sulfur atom S_α proximal to the peptide carrier and sulfur atom S_β proximal to the oligonucleotide. The disulfide was flanked on both sides by ethylene groups, and therefore S_α and S_β were expected to be closely similar both in intrinsic reactivity and steric terms.

Four substrate constructs, POC1 to POC4 (Table 2.1), were tested in which the length of the peptide carrier was adjusted two amino acid residues (Gly and Ser) at a time. When a substrate is fully stretched, each additional dipeptide lowers the disulfide

a further ~7 Å into the barrel. Seven single-cysteine αHL mutants were generated at positions 113, 115, 117, 119, 121, 123 and 143 (on one of the seven αHL subunits) in which the side chains point into the lumen of the transmembrane β barrel. Adjacent cysteines along a β strand are separated by an average vertical distance of ~5.6 Å (Cα-Cα), whereas Cys-143 on the antiparallel strand is ~4.6 Å (Cα-Cα) away from Cys-115. Given the experimental and computational evidence for a collinear trisulfide-like transition state in thiol-disulfide interchange (*25*, *26*, *49*), we reasoned that proximity and longitudinal alignment of the substrate disulfide with the nanoreactor cysteine side chain would be critical in achieving site-selective and regioselective control of the reaction.

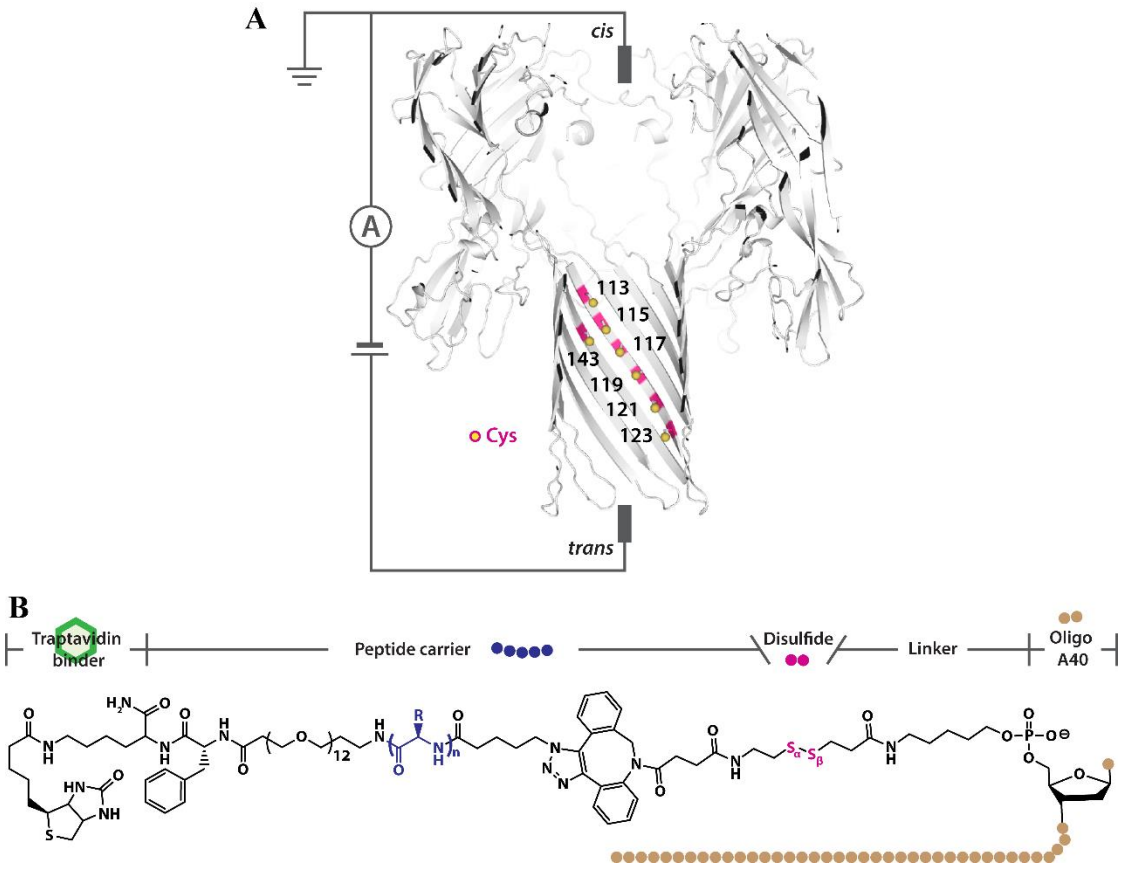


Fig. 2.1. Thiol-disulfide interchange at single-molecule level. (**A**) Cysteines (pink) on one of the seven α-hemolysin (αHL) subunits provide nucleophilic thiolates at defined inter-sulfur distances on a β-barrel scaffold. Six single-cysteine mutants (113, 115, 117, 119, 121, 143) and two double-cysteine mutants (115/117 and 115/143) were used. (**B**)

Substrates consisted of a biotin at one terminus followed by a peptide-oligonucleotide junction formed by click chemistry. The cleavable disulfide bond was installed between the click product and the oligo. The peptide length was manipulated by adding two amino acids (glycine or serine, Δn = 2) at a time following the PEG unit: POC1 (n = 0); POC2 (n = 2), POC3 (n = 4), POC4 (n = 6).

2.2 Turnover of thiol-disulfide interchanges

To monitor the site-selectivity and regioselectivity of thiol-disulfide interchange, a means of substrate turnover was first established. When capped by traptavidin, the substrate entered the αHL pore from the cis compartment under an applied positive potential and was stretched within the electric field. Subsequent alignment-mediated thiol-disulfide interchange with the nanoreactor cysteine side chain resulted in the formation of a covalent adduct within the nanoreactor (Fig. 2.2), which was revealed by a step in the ionic current passing through the pore. The adduct was released from the nanoreactor by the reducing reagent, 1,4-dithiothreitol (DTT) present in the trans compartment, which thereby allowed a new cycle to begin.

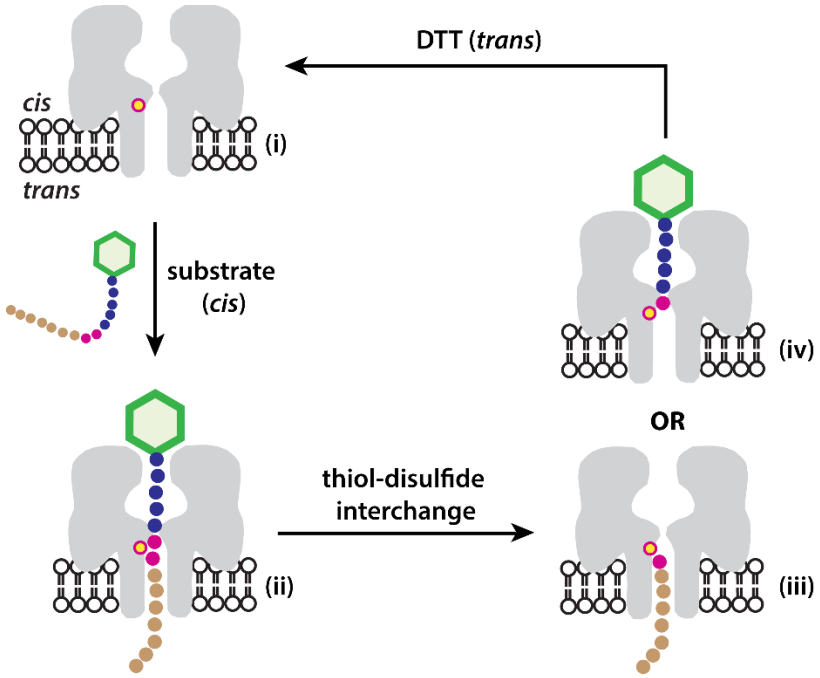


Fig. 2.2. Turnover of thiol-disulfide interchanges. A molecule of traptavidin-capped substrate threads into the nanopore from the cis side under an applied potential (+150 mV, trans), and translocation is arrested by the traptavidin stopper (ii). Proper alignment enables thiol-disulfide interchange between the cysteine thiolate and the substrate disulfide, generating either an αHL-peptide adduct (top) or an αHL-oligo adduct (bottom) (iii). Reduction of the newly formed disulfide by DTT (5 mM in the trans compartment) returns the nanopore to its open state (i) to initiate the next reaction cycle.

2.3 Controlled reaction between the nanoreactor thiolate and the substrate disulfide

In previous computational studies, thiol-disulfide interchange has been proposed to proceed through an almost linear trisulfide-like transition state with adjacent sulfur atoms separated by ~2.5 Å (*25*, *26*). Assuming free rotation of the nanoreactor cysteine side chains, the vertical distance (Fig. 2.3A) covered by a nucleophilic thiolate is ~3 Å (C-S bond length ~1.8 Å and C-C-S bond angle ~109.5°). For interchange to proceed, the 3 Å region must overlap with a 2.5 Å reactive zone at one end or the other of a disulfide (Fig. 2.3). When the transition state arrangement is inaccessible, with the nucleophilic thiolate residing outside the 2.5 Å radius at one end of a disulfide, the interchange is considered unfavourable (Fig. 2.3).

The system reported here can deliver a disulfide bond to a nucleophilic thiolate with angstrom precision. When the alignment was favourable, the interchange between the substrate disulfide and the nanoreactor thiolate generated a covalent adduct—either an αHL-peptide adduct or an αHL-oligo adduct (discussed later)—which was subsequently released by DTT (Fig. 2.2 and Fig. 2.4). When the alignment was unfavourable, the DTT (5 mM) diffusing from the trans compartment cleaved the substrate disulfide before interchange could occur and the resulting two fragments exited the nanopore, the DNA propelled by the applied potential and the peptide

segment by diffusion (Fig. 2.4). Based on these observations, reactivity was defined—regardless of regioselectivity—as the number of thiol-disulfide interchange reactions observed under specified conditions divided by the total number of substrate threading events (reactivity = n_{react}/n_{thread}) (Fig. 2.3C).

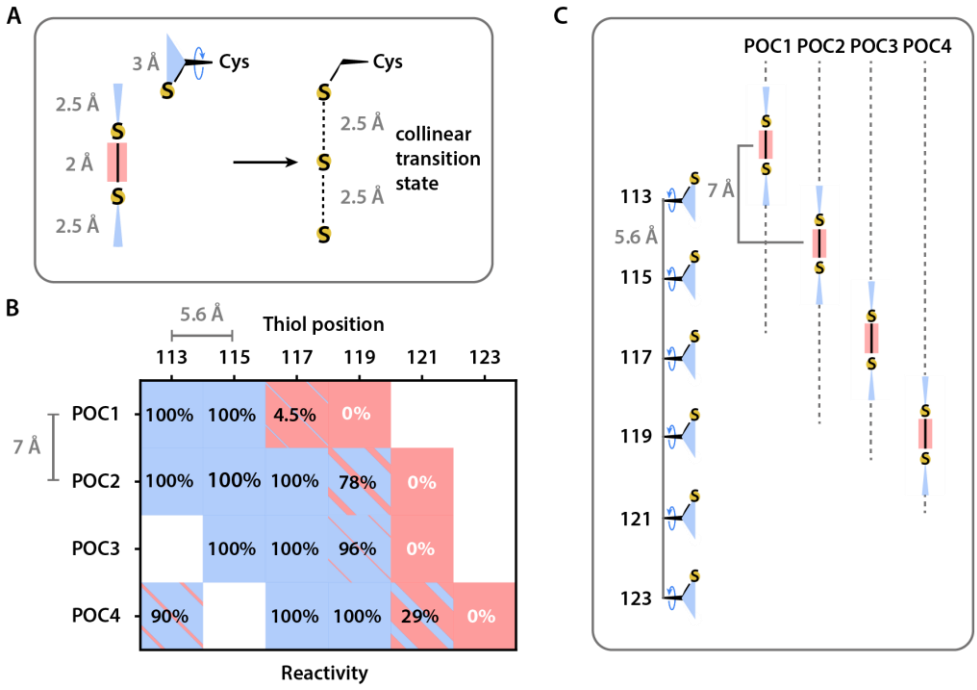


Fig. 2.3. Alignment of a substrate disulfide with a nanoreactor cysteine sulfur determines reactivity. (**A**) The computed transition state for a thiol-disulfide interchange has a S-S bond length of ~2.5 Å and a S-S-S bond angle of ~180°. This creates a reactive zone of 2.5 Å at either end of a disulfide (shown as arbitrary 10° blue wedges) and a forbidden zone of 2 Å in the middle (red). Assuming free rotation of the cysteine (Cys) side chains, a thiolate can access 3 Å in the vertical dimension (blue triangle). Overlap between the disulfide reactive zone and the thiolate-accessible region leads to thiol-disulfide interchange. (**B**) The reactivity between different substrates and single-cysteine nanoreactors, represented by the ratio of thiol-disulfide interchange events to the total number of threading incidents (reactivity = $n_{interchange}/n_{thread}$) expressed as a percentage. Blue indicates that threading led to adduct formation. Red indicates that the substrate was directly cleaved by DTT (5 mM, trans) without adduct

formation. All threading events resulted in either adduct formation or cleavage. White indicates an untested combination. (**C**) The relative positions of the substrates and the cysteines within single-cysteine nanoreactors. Accessible positions of the disulfide are indicated with dashed lines, which are inferred from the reactivity results. Successive nanoreactor cysteines are separated by ~5.6 Å vertically, due to the inclination of the β strand with respect to the barrel axis. The disulfide in POC1 was aligned based on based on its reactivity with Cys-113 and Cys-115. Assuming the force produced by the electric field is enough to fully elongate the substrates, two additional amino acids create a ~7 Å shift in disulfide depth inside the barrel, in a static model, and the disulfides in POC2, POC3 and POC4 were subsequently aligned on this basis.

Using the aforementioned four substrates (POC1 to POC4) and seven nanoreactors (nanoreactors 113, 115, 117, 119, 121, 123, and 143), we achieved a broad range of reactivity from 0% to 100% (Table 2.1, Fig. 2.3C). Notably, extension of the peptide carrier by two amino acids (~7 Å) dramatically affected reactivity. For example, once threaded into nanoreactor 117, the disulfide in POC1 was almost solely cleaved by DTT (4.5% reactivity), which suggests unfavourable alignment. In contrast, the disulfide in POC2, two amino acids longer, always formed a covalent adduct with nanoreactor 117 (100% reactivity) (Table 2.1, Fig. 2.3C). Similar phenomena were observed with nanoreactor 119 (POC1, 0% reactivity; POC2, 78%) and nanoreactor 121 (POC3, 0% reactivity; POC4, 29%). Likewise, shifting the cysteine location by two residues along the β strand (~5.6 Å vertically) was sufficient to greatly alter or prevent reaction with substrates. For example, when nanoreactor 119 is compared with nanoreactor 121, reactivity drops from 78% to 0% with POC2, from 96% to 0% with POC3, and from 100% to 29% with POC4.

We thereby demonstrate that, by manipulating the disulfide position within the substrate or the cysteine location within the nanoreactor, we can control whether or not thiol-disulfide interchange proceeds without altering the associated chemistry.

Table 2.1. Reactivity, selectivity and kinetics of thiol-disulfide interchange reactions

Substrate[a]	Peptide Carrier[b]	Nano-reactor	Reactivity	Regioselectivity[c]		k_β / s^{-1}[d]	k_α / s^{-1}[d]	k_{DTT} / s^{-1}[d]
				S_β	S_α			
POC1	bKF(PEG12)	113	100%	78%	22%	0.077 ± 0.011	0.023 ± 0.006	/[e]
		115	100%	100%	0%	0.69 ± 0.11	/	/
		117	4.5% (12/264)	100%	0%	0.00053 ± 0.00031	/	0.011 ± 0.001
		119	0%	/	/	/	/	0.016 ± 0.004
		143	100%	100%	0%	0.11 ± 0.01	/	/
POC2	bKF(PEG12) GG	113	100%	44%	56%	0.028 ± 0.006	0.036 ± 0.007	/
		115	100%	79%	21%	1.7 ± 0.3	0.44 ± 0.14	/
		117	100%	89%	11%	0.055 ± 0.005	0.0069 ± 0.0018	/
		119	78% (72/92)	100%	0%	0.032 ± 0.004	/	0.0090 ± 0.0020
		121	0%	/	/	/	/	0.014 ± 0.003
		143	100%	90%	10%	0.45 ± 0.06	0.05 ± 0.02	/
POC3	bKF(PEG12) SGGS	115	100%	59%	41%	1.8 ± 0.3	1.2 ± 0.3	/
		117	100%	70%	30%	0.17 ± 0.03	0.079 ± 0.019	/
		119	96% (55/57)	95%	5%	0.15 ± 0.01	0.0084 ± 0.0050	/[f]
		121	0%	/	/	/	/	0.0092 ± 0.0009
POC4	bKF(PEG12) SGGSGG	113	90% (37/41)	19%	81%	0.019 ± 0.007	0.082 ± 0.012	/[f]
		117	100%	35%	65%	0.094 ± 0.017	0.17 ±0.02	/
		119	100%	83%	17%	0.36 ± 0.05	0.07 ±0.02	/
		121	29% (23/78)	100%	0%	0.0038 ± 0.0008	/	0.0091 ± 0.0007
		123	0%	/	/	/	/	0.018 ± 0.004

[a] An oligoadenosine 40mer (A40) was used in all substrates.
[b] bK, biotinyl-L-lysine; F, phenylalanine; PEG, polyethylene glycol; S, serine; G, glycine.
[c] Regioselectivity is presented as the percentage of interchanges resulting in an αHL-oligo adduct (S_β) or an αHL-peptide adduct (S_α). DTT cleavages are discounted.
[d] Dwell time analysis and rate constant estimations were performed by using the maximum interval likelihood algorithm of QuB.
[e] /, not measurable.
[f] The DTT cleavage occurred < 5 times hence was disregarded in the kinetic model.

2.4 Sulfur atom selection within the substrate disulfide

The substrate disulfide connects a neutral segment (the peptide) with a negatively charged polymer (the oligonucleotide) (Fig. 2.1B). Upon thiol-disulfide interchange with a cysteine nanoreactor, either an αHL-peptide adduct or an αHL-oligo adduct is formed depending on whether the S_α or the S_β atom of the disulfide bond is attacked (Fig. 2.4). The two adducts were discriminated by their distinctive current signatures under an applied potential (Table 2.2, Fig. 2.4). In contrast to the stable current levels associated with αHL-oligo adducts (S_β attacked), rapid switching between three discrete levels was observed for αHL-peptide adducts (S_α attacked).

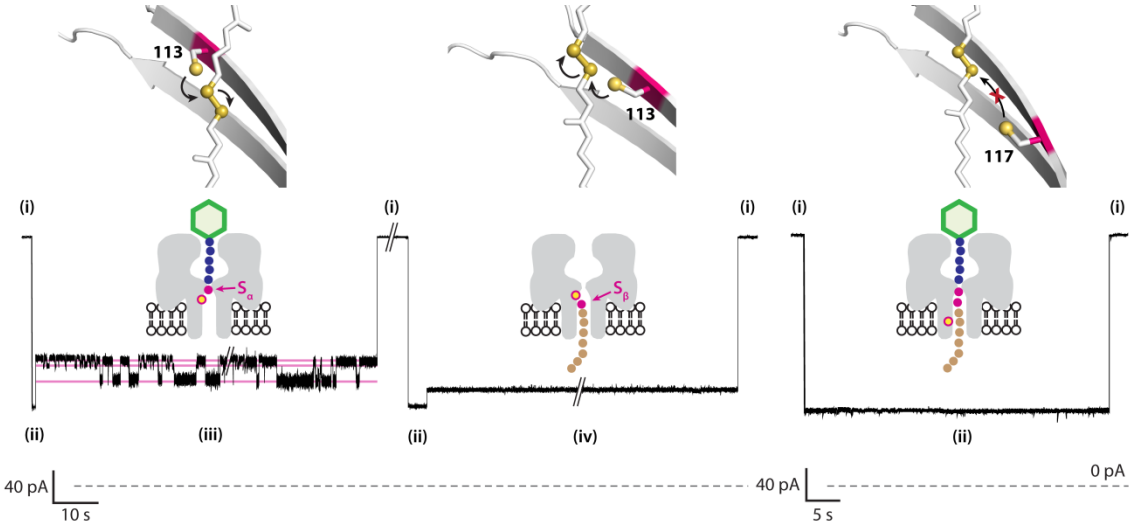


Fig. 2.4. Identification of turnover intermediates from current signatures. Threading of POC1 into nanoreactor 113 led to partial blockade of the ion flow ($I_{res\%}$ = 33%) (ii). Thiol-disulfide interchange between nanoreactor 113 and POC1 produced two types of electrical patterns (iii, iv) corresponding to the formation of two different covalent adducts with the nanopore: the αHL-peptide adduct offered a number of sub-conductance states (pink lines) (iii) ($I_{res\%}$ = 43%, 48%, 51%); the αHL-oligo adduct gave a single current state (iv) ($I_{res\%}$ = 39%). Subsequent adduct release by DTT returned the system to its initial state (i). Misalignment of the substrate disulfide and the thiolate disfavoured the reaction between nanoreactor 117 and POC1, resulting mainly in direct cleavage of the substrate disulfide by DTT rather than adduct formation (4.5% reactivity). Traces were filtered at 200 Hz. Conditions: 2 M KCl, 20 mM HEPBS, 20 μM EDTA, 5 mM DTT (trans), pH 8.5, 20 ± 1 °C.

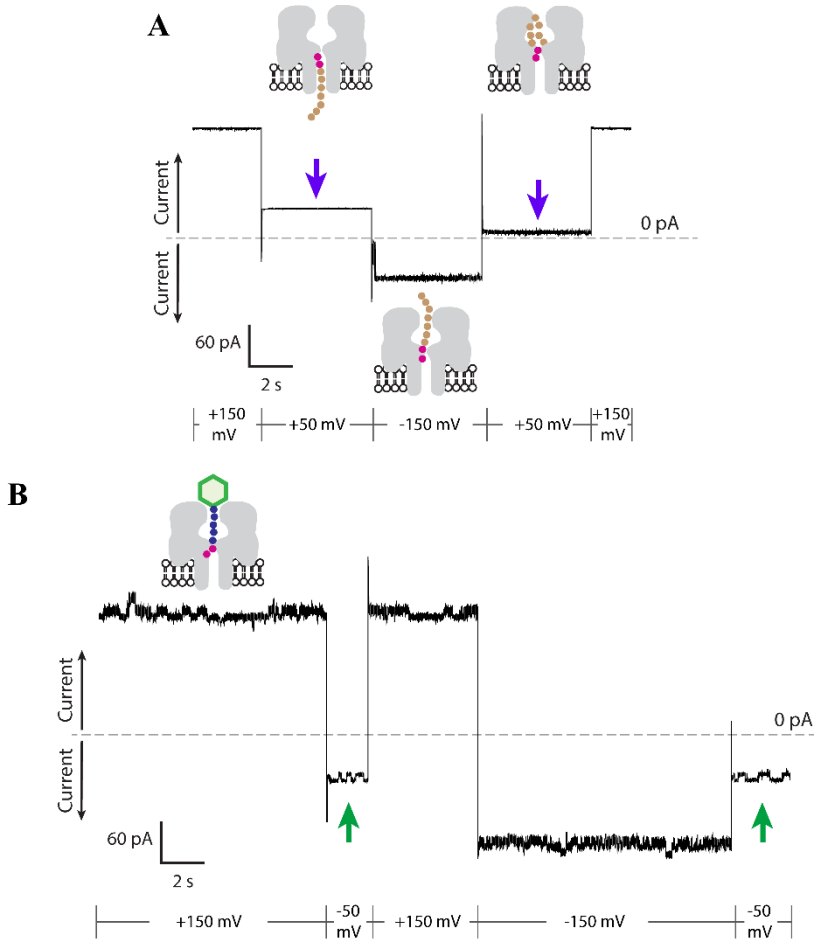


Fig. 2.5. Current changes caused by αHL-polymer adducts in response to voltage steps. (**A**) An αHL-oligo adduct was formed after thiol-disulfide interchange between POC2 and nanoreactor 119. A step to +50 mV produced a different level of current blockade depending on whether the step was made after -150 mV or after +150 mV (purple arrows), indicating that a switch in oligonucleotide orientation occurs at high applied potentials. (**B**) An αHL-peptide adduct was formed after thiol-disulfide interchange between POC2 and nanoreactor 113. After a step to -50 mV, the current levels of the three sub-conductance states were the same whether the step came after +150 mV or after -150 mV (green arrows), in keeping with a covalent adduct locked inside the vestibule of the αHL pore by the traptavidin stopper. Traces were filtered at 200 Hz. Conditions: 2 M KCl, 20 mM HEPBS, 20 μM EDTA, 5 mM DTT (trans), pH 8.5, 20 ± 1 °C.

We speculate that, within the electric field inside the αHL barrel (*56*), the charged oligonucleotide experiences a constant pulling force that locks its conformation in a stretched state, whereas the neutral peptide is less constrained, switching between

several conformations. To confirm the identity of adducts formed upon thiol-disulfide interchange, we also probed their behaviour in response to voltage steps (Fig. 2.5). For an αHL-oligo adduct (Fig. 2.5A), the negative charges on the oligonucleotide within the barrel re-orientated the chain upon switching the applied potential between high positive and negative values (e.g. ±150 mV). At low potentials (e.g. +50 mV), the pulling force was insufficient to overcome the energy barrier for re-orientation. Therefore, due to condensation of the oligonucleotide, the current level recorded at a low potential (e.g. +50 mV) was lower if the immediately previous applied potential was of an opposite polarity. For an αHL-peptide adduct (Fig. 2.5B), the peptide fragment was locked inside the upper vestibule of the αHL pore with traptavidin acting as a stopper. In this case, the current pattern at a low applied potential (e.g. -50 mV) was the same following high potentials of opposite polarity (e.g. ±150 mV), which means that the conformational flexibility was not affected.

By counting adducts formed over multiple substrate turnovers, we determined the sulfur regioselectivity for each nanoreactor-substrate combination, presented as the percentage of interchanges that resulted in attack on S_α or S_β (Fig. 2.6A, Table 2.1). In general, when a cysteine was located within the reactive zone of a substrate disulfide, lowering its position (e.g. from 113 to 115) reduced the chance of forming an αHL-peptide adduct and increased the chance of forming an αHL-oligo adduct. Complete regioselectivity towards S_β (αHL-oligo adduct formation) was achieved with POC1 and nanoreactor 115 or 143, POC2 and nanoreactor 119, and POC4 and nanoreactor 121 (Fig. 2.6). In these cases, the nucleophilic thiolate could only reach the lower 2.5 Å reactive zone of the disulfide (Fig. 2.2A), permitting formation only of the transition state en route to the αHL-oligo adduct.

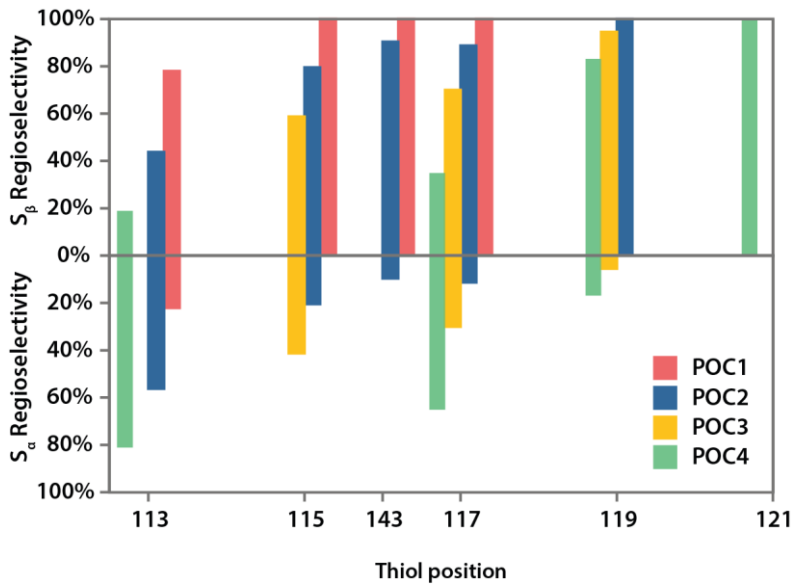


Fig. 2.6. Sulfur regioselectivity of thiol-disulfide interchanges. The percentage of interchanges that resulted in formation of either an αHL-peptide adduct or an αHL-oligo adduct for thiol-disulfide interchanges between substrates POC1 (red), POC2 (blue), POC3 (yellow), and POC4 (green) and nanoreactors 113, 115, 143, 117, 119, and 121.

The completely S_β regioselective thiol-disulfide interchange between POC1 and nanoreactor 115 suggests that the thiolate resides below S_β in the barrel. Given that the fully stretched linker (Fig. 2.1B) extends 13 Å and that Cys-115 is located ~40 Å from the exit of the barrel, ~4 phosphodiester bonds of POC1 must be located within the barrel's electric field. Each extension of the peptide carrier by two amino acids moves one phosphodiester bond out of the field (the inter-nucleotide distance in stretched single-stranded DNA is 6.9 Å (*75*)). The controlled positioning of a disulfide requires at least one charged residue inside the barrel to execute the pulling, which makes POC4 the longest workable construct in the present set-up. Over 80% regioselectivity towards S_α (αHL-peptide adduct formation) was achieved with POC4 and nanoreactor 113 (Fig. 2.6). We envisage that complete S_α regioselectivity can be achieved by using a longer peptide carrier and a longer pore.

Table 2.2. Residual current levels ($I_{res\%}$) for individual substrate/nanoreactor combination

+150 mV[a]	**$I_{res\%}$**		
POC1	**Threading[b]**	**αHL-Oligo[b]**	**αHL-Peptide[c]**
113	33%	39%	43%-51%
115	31%	45%	/
117	34%	47%	/
143	31%	44%	/
115/117	30%	45%, 47%	/
115/143	30%	45%	/
POC2	**Threading**	**αHL-Oligo**	**αHL-Peptide**
113	33%	39%	39%-56%
115	32%	45%	27%-50%
117	33%	47%	42%-49%
119	34%	53%	/
143	32%	44%	27%-54%
POC3	**Threading**	**αHL-Oligo**	**αHL-Peptide**
115	32%, 33%, 34%[d]	45%	24%-52%
117	32%, 33%, 34%[d]	47%	29%-56%
119	32%, 33%, 34%[d]	53%	28%-48%
121	32%, 33%, 34%[d]	/	/
POC4	**Threading**	**αHL-Oligo**	**αHL-Peptide**
113	33%, 34%, 35%[d]	39%	28%-56%
117	33%, 34%, 35%[d]	47%	29%-54%
119	33%, 34%, 35%[d]	53%	33%-54%
121	33%, 34%, 35%[d]	/	/
123	34%, 35%, 36%[d]	/	/

[a] All the measurements in the table were carried out at +150 mV.

[b] The standard deviations of $I_{res\%}$ are less than 0.5% (derived from n > 3 separate experiments for each substrate/nanoreactor combination).

[c] Three discrete interchanging levels were recorded for each adduct. However, adduct to adduct variation in $I_{res\%}$ was seen. The observed range of $I_{res\%}$ is reported here.

[d] Three discrete interchanging levels were recorded at the threading stage.

2.5 Kinetics of thiol-disulfide interchange

After a substrate threads into a nanoreactor, one of three possible reactions occurs: generation of an αHL-peptide adduct, generation of an αHL-oligo adduct, or cleavage by DTT. At the single molecule level, rate constants are related to the mean dwell times of states (*55*). For a unimolecular adduct formation, the rate constant—k_α for the formation of αHL-peptide adduct; k_β for the formation of αHL-oligo adduct—is the reciprocal of the corresponding mean dwell time τ_α or τ_β: $\langle\tau_{\alpha,\beta}\rangle = 1/ k_{\alpha,\beta}$. For the bimolecular DTT cleavage at a fixed DTT concentration (e.g. 5 mM), $\langle\tau_{DTT}\rangle = 1/ k'_{DTT} = 1/ k_{DTT} [RS^-]_{DTT}$ (k'_{DTT} is the pseudo-first order rate constant dependent on $[RS^-]_{DTT}$, the concentration of the DTT thiolate).

Over multiple turnovers, the mean waiting time before reaction ($\langle\tau\rangle$) is:

$$\langle\tau\rangle = \frac{1}{k_\alpha + k_\beta + k'_{DTT}} \qquad \text{Eq. 2.1}$$

$$\langle\tau\rangle = \frac{\langle\tau_\alpha\rangle\langle\tau_\beta\rangle\langle\tau_{DTT}\rangle}{\langle\tau_\alpha\rangle\langle\tau_\beta\rangle + \langle\tau_\beta\rangle\langle\tau_{DTT}\rangle + \langle\tau_\alpha\rangle\langle\tau_{DTT}\rangle} \qquad \text{Eq. 2.2}$$

Depending on the alignment, one of the three pathways may dominate, e.g. $\langle\tau_\alpha\rangle$, $\langle\tau_{DTT}\rangle \gg \langle\tau_\beta\rangle$, in which case $\langle\tau\rangle = \langle\tau_\beta\rangle$. In cases where more than one pathways compete, dwell time analysis was performed to derive the independent rate constants (k_α, k_β or k_{DTT}') by using the maximum interval likelihood (MIL) algorithm of QuB software (*76*) (Table 2.1, Fig. 2.7).

Given that the thiol-disulfide interchange involves the thiolate anion, the rate constants k_α and k_β depend on not only the alignment, but also on the pH and the pK_a value of the nanoreactor thiol. The latter is estimated to be 8.5 ± 1.5 in a wide β barrel (*8*). Nevertheless, when comparing different substrates reacting with the same nanoreactor (i.e. pK_a is fixed) at a fixed pH (e.g. pH 8.5), k_α and k_β directly correlate with the thiolate alignment relative to S_α and S_β, respectively. The spatial relationship between substrates

and nanoreactors can therefore be inferred from the derived rate constants (Table 2.1, Fig. 2.7).

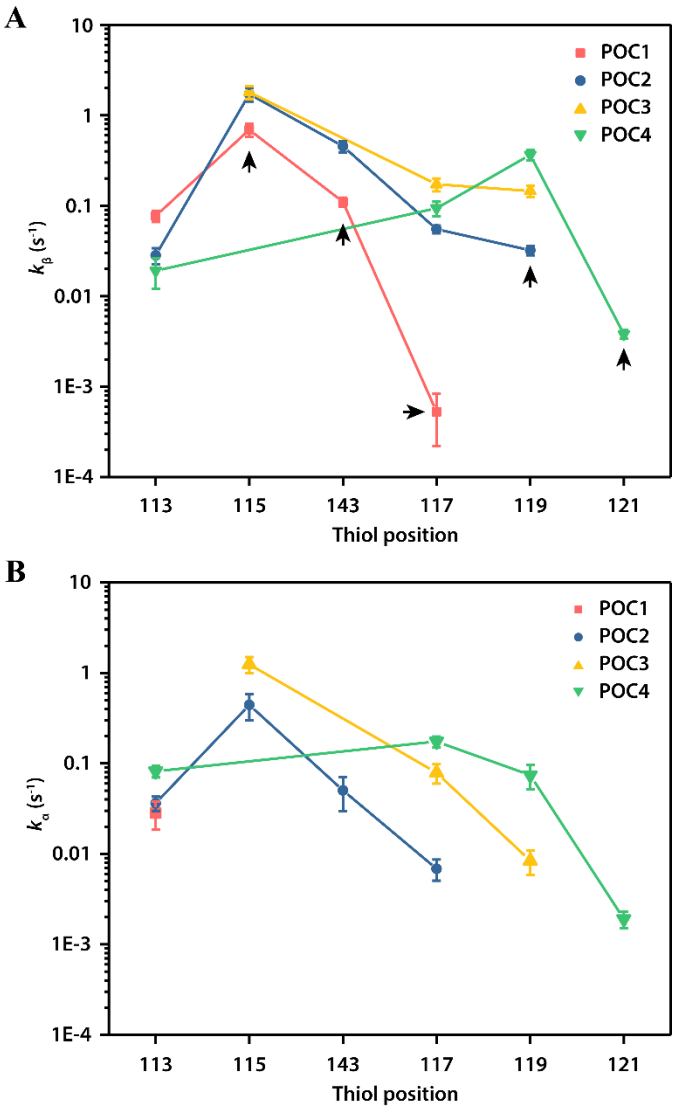


Fig. 2.7. Kinetics of single-molecule thiol-disulfide interchange. (**A**) The rate constant of αHL-oligo adduct formation (k_β) for each tested combination of substrates and nanoreactors. Arrows indicate the cases with complete S_β regioselectivity. (**B**) The rate constant of αHL-peptide adduct formation (k_α) for each tested combination of substrates and nanoreactors. Dwell time analysis and rate constant estimations were performed by using the maximum interval likelihood (MIL) algorithm of QuB.

For each nanoreactor, optimal alignment with the substrate S_β atom was found based on the rate of αHL-oligo adduct formation (k_β). The highest k_β values corresponds to the most favourable alignment achieved (Fig. 2.3C), e.g. between nanoreactor 113

and POC1, nanoreactor 115 and POC3, nanoreactor 117 and POC3, as well as nanoreactor 119 with POC4 (Fig. 2.7A). A ~300-fold difference in k_β values was seen with nanoreactor 117, which gave 5.3×10^{-4} s^{-1} with POC1 and 0.17 s^{-1} with POC3. This shifts the overall outcome from domination by DTT cleavage with POC1 ($k_{DTT} = 9.4 \times 10^{-3}$ s^{-1}) to domination by thiol-disulfide interchange at S_β with POC3, showcasing the influence of alignment on reaction kinetics and selectivity. Interestingly, we found that complete S_β regioselectivity was always achieved with the smallest k_β among tested substrates (Fig. 2.6 and Fig. 2.7A). This implied that regioselectivity can be improved at the expense of reaction rate in a dynamic system (discussed later).

2.6 Site selective and regioselective tandem thiol-disulfide interchanges

To recapitulate the tandem thiol-disulfide interchanges frequently seen with dithiol proteins (*4*, *77*), we constructed two double-cysteine nanoreactors based on nanoreactor 115 with additional cysteines on either the antiparallel strand (115/143) or on the same strand (115/117) within a single aHL subunit. Using POC1—which reacted with the three single-cysteine nanoreactors (115, 143, and 117) with complete S_β regioselectivity—we demonstrated that the double-cysteine nanoreactors undergo two consecutive thiol-disulfide interchanges.

After a POC1 molecule threaded into a double-cysteine nanoreactor, the substrate disulfide first reacted with one of the two available cysteine thiolates to form a covalent αHL-oligo adduct. Site-selectivity of the first thiol-disulfide interchange is reflected by the chance of adduct formation at either cysteine site. With nanoreactor 115/117, oligo-adducts at positions 115 and 117 can be differentiated by their residual currents ($\Delta Ires\% = Ires\%_{(117)} - Ires\%_{(115)} = 2.0\% \pm 0.5\%$, n = 3) (Fig. 2.8A), which

match the results obtained using single-cysteine nanoreactors ($Ires\%_{(115)}$ = 45% ± 1%, n = 9; $Ires\%_{(117)}$ = 47% ± 1%, n = 12). With nanoreactor 115/143, αHL-oligo adducts formed at positions 115 and 143 can also be differentiated by their residual currents ($\Delta Ires\% = Ires\%_{(115)} - Ires\%_{(143)}$ = 0.68% ± 0.03%, n = 3) (Fig. 2.8B and Fig. 2.9), which is consistent with the results obtained with single-cysteine nanoreactor ($Ires\%_{(115)}$ = 45% ± 1%, n = 9; $Ires\%_{(143)}$ = 46% ± 1%, n = 6). With nanoreactor 115/117, we found that the substrate disulfide reacted site-selectively (100%) with Cys-115 to form the αHL-oligo adduct (n = 30), which agrees with the kinetics of the single-cysteine nanoreactors: $k_{\beta(115)} \gg k_{\beta(117)}$. In contrast, we observed 81% site-selectivity towards Cys-115 and 19% site-selectivity towards Cys-143 with nanoreactor 115/143 (Fig. 2.8B and Fig. 2.9).

Once the adduct had been formed, the free cysteine on nanoreactor 115/117 or nanoreactor 115/143 initiated a second thiol-disulfide interchange by attacking either the sulfur contributed by the nanoreactor cysteine (Sγ) or the sulfur contributed by the oligonucleotide (S_β) (Fig. 2.8). Surprisingly, the regioselectivities of these two intramolecular reactions differed. When the free cysteine was on the same strand (115/117), complete S_β regioselectivity resulted in transfer of the oligonucleotide back and forth from one site to another (Fig. 2.8B). In the presence of DTT, the average number of transfers before adduct release by DTT (Fig. 2.8A) was 11 over 40 cycles and up to 29 successive transfers have been recorded in a single cycle. When the free cysteine was on the antiparallel strand (115/143), we observed thiolate attack on both S_β and Sγ (see below), the latter of which led to the irreversible release of the oligonucleotide from the nanoreactor and formation of a cross-strand disulfide in the absence of DTT (Fig. 2.8B). Subsequent addition of DTT regenerated free cysteines through a transient αHL-DTT intermediate (*58*) (Fig. 2.8B and Fig. 2.9). Depending on which site an αHL-oligo adduct initially formed on, the average number of transfers

before intramolecular release varied: 0.54 transfers over 48 cycles for the αHL-oligo adducts initially formed on Cys-115 and 1.36 over 11 cycles for those initially formed on Cys-143. Up to 6 transfers have been seen within a single cycle (Fig. 2.9). Among the αHL-oligo adducts formed on Cys-115, 85% was immediately released through intramolecular Sγ attack by Cys-143 without transfers. However, all the αHL-oligo adducts formed on Cys-143 had been transferred at least once to Cys-115 before release occurred.

In summary, with nanoreactor 115/117, the first thiol-disulfide interchange was site-selective towards Cys-115 and S_β regioselective; the second interchange showed complete S_β regioselectivity leading to transfer only (no release). With nanoreactor 115/143, the first thiol-disulfide interchange was S_β regioselective and 81% site selective towards Cys-115; the second interchange exhibited 85% Sγ regioselectivity on Cys-115 leading to adduct release. No release was observed on Cys-143. The observed site selectivities are in agreement with the single-cysteine kinetic results (Table 2.1). The regioselectivities of the second thiol-disulfide interchange were attributed to the Cα-Cα distances between the two cysteines involved (Fig. 2.8 and Fig. 2.9). For nanoreactor 115/117, only the collinear alignment along the β strand (Cα-Cα = 6.8 Å) was productive resulting in S_β attack (Fig. 2.8). For nanoreactor 115/143, two collinear arrangements were possible across the β strands (Cα-Cα = 4.6 Å) of which thiolate attack on Sγ was more favourable on Cys-115 (Fig. 2.8 and Fig. 2.9). Future computational modelling may unravel the conformational energetics leading to the observed biased regioselectivity under an applied potential.

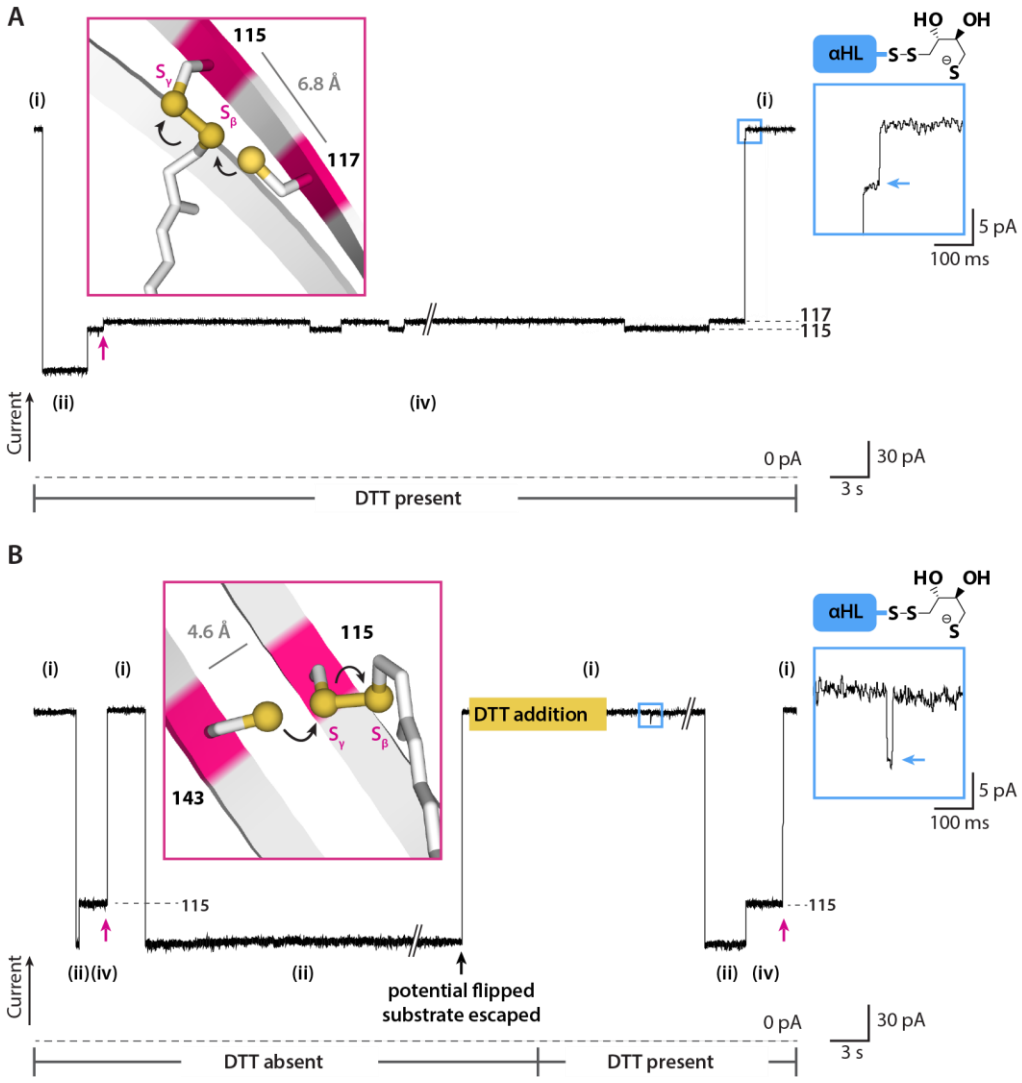


Fig. 2.8. Tandem thiol-disulfide interchanges with double-cysteine nanoreactors. (**A**) With nanoreactor 115/117, POC1 reacted regioselectively and site-selectively with cysteine 115 to form an αHL-oligo adduct. Subsequent regioselective attack on S_β initiated intramolecular thiol-disulfide interchange to transfer the adduct to cysteine 117 (pink arrow). The adduct moved back and forth between the two sites on the same β strand before it was eventually released by DTT, which was evidenced by an intermediate αHL-DTT adduct (blue arrow). (**B**) With nanoreactor 115/143, POC1 reacted 81% site-selectively with cysteine 115 to form an αHL-oligo adduct. Attack on S_γ by Cys-143 on the antiparallel strand cleaved the adduct from the nanoreactor (pink arrow). Consequently, a cross-strand disulfide bridge was generated, prohibiting turnover of a new substrate until DTT regenerated the two free cysteines through an intermediate αHL-DTT adduct (blue arrow). Traces were filtered at 200 Hz. Conditions: 2 M KCl, 20 mM HEPBS, 20 µM EDTA, 5 mM DTT (trans), pH 8.5, 20 ± 1 °C.

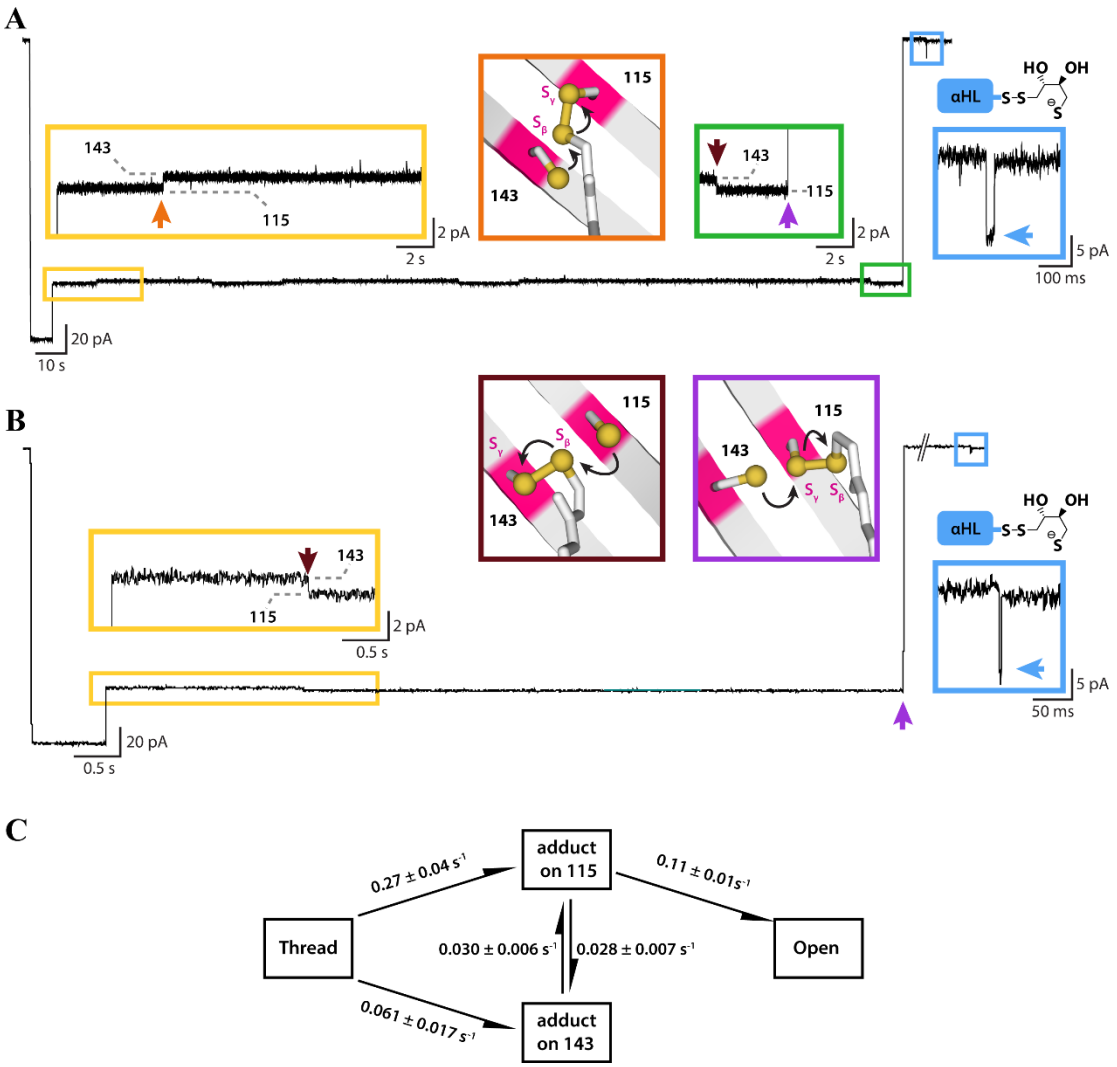


Fig. 2.9. Thiol-disulfide interchanges between nanoreactor 115/143 and POC1. (**A**) An αHL-oligo adduct was formed at site 115 following threading. The oligonucleotide transferred between Cys-115 and Cys-143 via S_β attack (orange: transfer from 115 to 143; brown: transfer from 143 to 115) until S_γ is attacked to form the cross-strand disulfide (purple). An αHL-DTT adduct was subsequently observed (blue) as DTT reduced the disulfide to regenerate the two free cysteines. (**B**) An αHL-oligo adduct was formed at site 143 following threading. The oligonucleotide transferred to Cys-115 (brown) where the release took place through S_γ attack by Cys-143 (purple). Traces were filtered at 50 Hz. Insets showing αHL-DTT adducts were filtered at 1000 Hz. Conditions: 2 M KCl, 20 mM HEPBS, 20 µM EDTA, 5 mM DTT (trans), pH 8.5, 20 ± 1 °C. (**C**) The kinetic model and the derived rate constants for the consecutive thiol-disulfide interchanges observed with nanoreactor 115/143 and POC1. Dwell time analysis and rate constant estimations were performed by using the maximum interval likelihood algorithm of QuB.

2.7 Regioselectivity of thiol-disulfide interchange between DTT and αHL adducts

Previously, a transient intermediate was detected upon DTT reduction of a mixed-disulfide formed between Cys-117 and 5,5'-dithiobis-(2-nitrobenzoic acid) (DTNB) (*58*). In the present work, DTT was used to release covalent disulfide adducts—either αHL-oligo adducts or αHL-peptide adducts—from nanoreactors. The DTT thiolate can attack either the nanoreactor cysteine sulfur (S_γ) or the sulfur atom donated by the polymer: an oligo or a peptide (S_δ) (Fig. 2.10A). When S_γ is attacked, an αHL-DTT adduct is formed, which has a distinctive breakdown signal. We investigated the regioselectivity of release of the αHL-polymer adduct by DTT, by monitoring the formation of the αHL-DTT adduct (Table 2.3, Fig. 2.10A). The S_γ regioselectivity is defined as the number of αHL-DTT adducts detected against the overall number of events:

$$\text{regioselectivity} = \frac{n(S_\gamma)}{n(S_\gamma) + n(S_\delta)}$$

In general, DTT attacked S_δ more frequently than S_γ (Table 2.3). Complete S_δ regioselectivity was seen when the mixed disulfide was formed between the nanoreactor 113 or 143 and the polymer (an oligo or a peptide). In contrast, complete S_γ regioselectivity was seen when the mixed disulfide was formed between nanoreactor 113 and DTNB (Table 2.3, Fig. 2.10B), and over 80% reaction at S_γ was seen when the mixed disulfide was formed between nanoreactor 143 and DTNB. This ruled out the possibility that the αHL-DTT adducts were too short-lived to be detected.

For the thiol-disulfide interchange between DTT and covalent adducts, the constituent sulfurs are not chemically and sterically equivalent and the DTT is freely

diffusing in solution. Therefore, no spatial control could be exerted on the attacking thiolate and innate regioselectivity was observed (Table 2.3).

The lifetime of an αHL-DTT adduct (<$\tau_{\alpha HL\text{-}DTT}$>)—the mean waiting times before cyclization (Table 2.3)—varied between the nanoreactors. This implies that the nanoreactor cysteine thiols, which are the leaving group upon DTT thiolate attack, have varied pK_a.

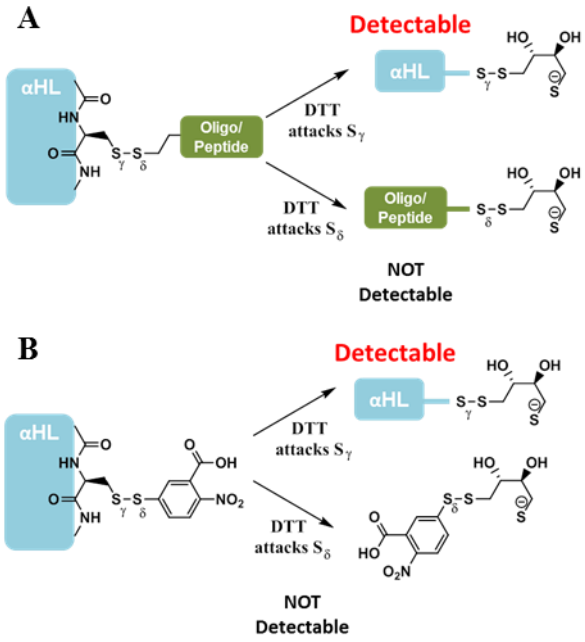


Fig. 2.10. Regioselectivity of thiol-disulfide interchange between DTT and αHL adducts. The mixed-disulfide is formed (**A**) with the macromolecular substrates to give either an αHL-oligo adduct or an αHL-peptide adduct, or (**B**) with DTNB (**B**). When the DTT attacks the cysteine sulfur (S_γ), the transient intermediate can be detected by single-channel electrical recording.

Table 2.3. Regioselectivity of thiol-disulfide interchange between DTT and αHL adducts, and the lifetimes of the αHL-DTT adducts

	DTT/αHL adducts (αHL-oligo adducts and αHL-peptide adducts combined)		**DTT/DTNB**	
	Regioselectivity $\frac{n(S_\gamma)}{n(S_\gamma) + n(S_\delta)}$	**$\langle\tau_{\alpha HL\text{-}DTT}\rangle$ / ms[a]**	**Regioselectivity** $\frac{n(S_\gamma)}{n(S_\gamma) + n(S_\delta)}$	**$\langle\tau_{\alpha HL\text{-}DTT}\rangle$ / ms[a]**
113	0/137 (0%)	/[b]	47/47 (100%)	85 ± 8
115	41/143 (29%)	220 ± 30	85/85 (100%)	150 ± 20
143	0/133 (0%)	/	82/100 (82%)	18 ± 2
117	59/220 (27%)	73 ± 6	63/63 (100%)	66 ± 6
119	3/198 (1.5%)	34 ± 30[c]	0/145 (0%)[d]	/
121	3/23 (13%)	110 ± 80[c]	0/60 (0%) [d]	/
	Detection of αHL-DTT adducts	**$\langle\tau_{\alpha HL\text{-}DTT}\rangle$ / ms[a]**	**Detection of αHL-DTT adducts**	**$\langle\tau_{\alpha HL\text{-}DTT}\rangle$ / ms[a]**
115/143	59/59 (100%)	9.4 ± 1.7	Not tested	

[a] Dwell-time analysis was performed by using the maximum interval likelihood algorithm of QuB unless stated otherwise.

[b] /, not measurable

[c] $\langle\tau_{DTT\text{-}\alpha HL}\rangle$ was calculated as the arithmetic mean due to insufficient events with nanoreactor 119 or 121.

[d] We speculate that the current levels of the adducts formed upon DTNB reaction with nanoreactor 119 or 121 cannot be distinguished from the current level of the αHL-DTT adduct. Future experiments with alternative reducing reagents (e.g. dihydrolipoic acid) might give identifiable current steps for the transient intermediates.

2.8 Discussions

In biology, thiol-disulfide interchange in proteins is highly site-selective and regioselective, which is achieved by controlling the accessibility or the chemical properties of the reactive thiols within an enzyme and its substrate (*4*, *78–80*). Until now, no artificial system has been constructed to mediate precise selective control over this important chemistry. In the present work, we examined the influence of spatial alignment on a thiol-disulfide interchange between a nanoreactor cysteine and a disulfide consisting of chemically equivalent sulfur atoms (S_α and S_β). Our results showed that thiol-disulfide interchange was sensitive towards alignment, which could be controlled with Angstrom precision. By manipulating the relative position of the substrate disulfide and the thiolate in single-cysteine nanoreactors, we steered the interchange between complete S_β regioselectivity and >80% S_α regioselectivity. Tandem thiol-disulfide interchanges with double-cysteine nanoreactors exhibited remarkable site-selectivity and regioselectivity simultaneously, which is present in biological systems but unprecedented in synthetic systems.

It is worth noting that our system is dynamic and we expect a complex distribution of relative positions of the disulfide and the thiolate (*81*). Assuming a static system, regioselectivity would be achieved with optimal alignment (i.e. fastest reaction rates). However, to achieve complete regioselectivity with a dynamic system which is presented here, all possible positions of the cysteine thiolate at equilibrium must overlap with the reactive zone of only one of the two sulfurs. This resulted in high regioselectivity achieved at less favourable alignment (i.e. slower reaction rates). This agrees with the obtained kinetic profile, where completely selective formation of adducts occurred at the slowest rate for a particular nanoreactor.

Previously, in the bottom-up fabrication of synthetic cells (*82*) and tissues (*83*), thiol-disulfide interchange-based redox signalling has seldom been used, presumably because it has been hard to control. Based on our studies, we envisage that cross-membrane communication based on regioselective alignment will allow: 1. cargo sorting to one side of the membrane or the other; 2. the controlled release of chemical signals; 3. the selective unmasking of thiol-functionalized materials for downstream signalling. Cross-membrane communication has been achieved with pore-forming proteins, such as αHL (*84*, *85*), or with synthetic membrane transducers or transmitters (*86–90*). The system reported here might be directly incorporated into droplet-based synthetic tissues scaffolded with lipid bilayers (*83*) to allow internal (inter-compartment) or external communication (Fig. 2.11).

In future work, the steric and electronic properties of functional groups flanking the disulfide in the substrate might be manipulated during synthesis to produce innate regioselectivity. It would be interesting to probe the influence of alignment on regioselectivity in such biased cases. The spatial alignment of reactants within a narrow 'tube' might be generalised for the site and regioselective control of other chemistries. Indeed, a click reaction with a dialkyne becomes site-selective when reactive azides are suitably spaced within the porous structure of a metal-organic framework (*71*). Our results will encourage future exploration of the spatial preorganisation to generate even finer differentiation between multiple sites within a functional group for regioselectivity modulation.

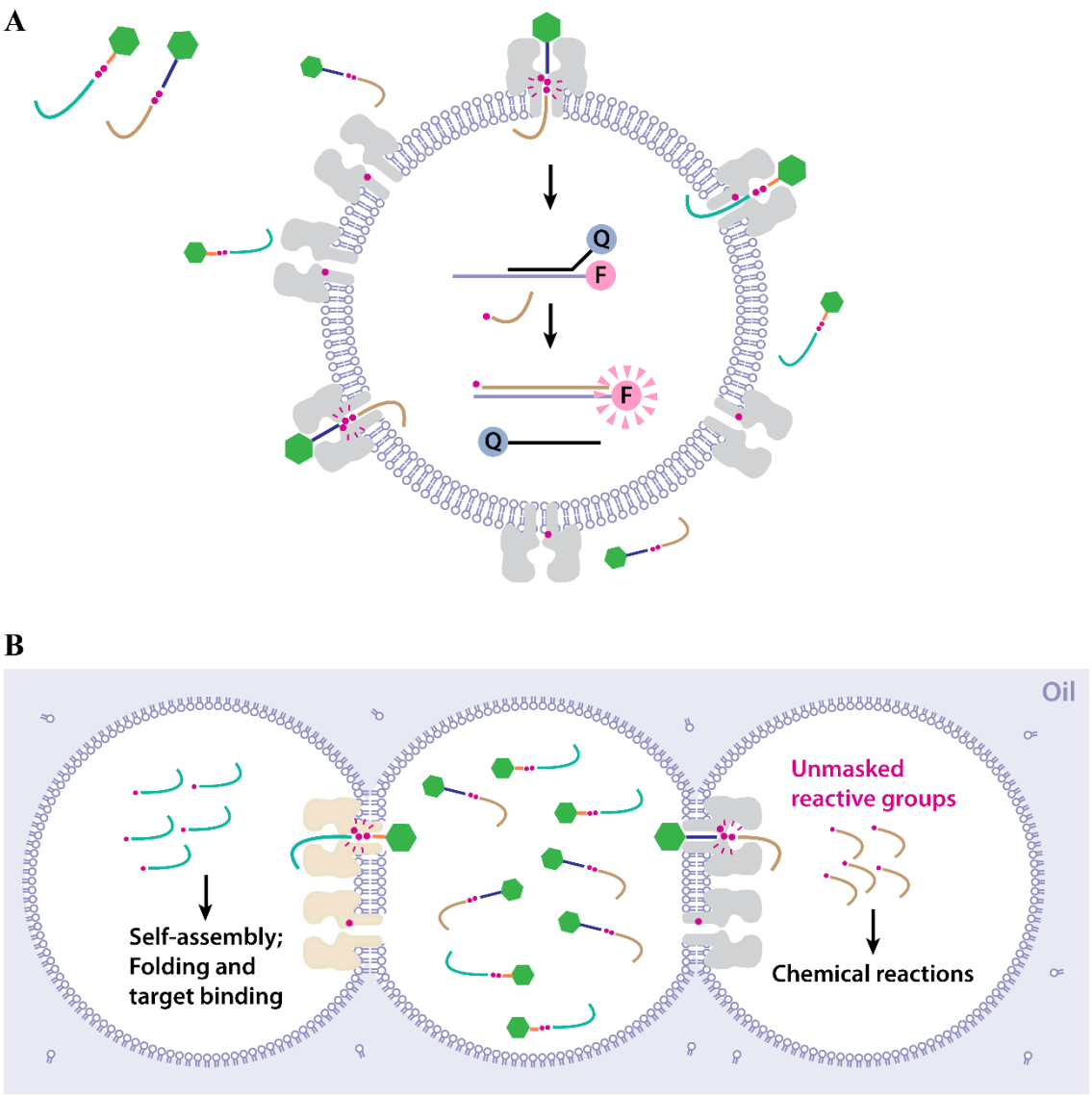


Fig. 2.11. Potential applications. (**A**) Among a mixture of substrates in solution, only the brown cargos are delivered into a giant unilamellar vesicle (GUV) by selective reactions between the membrane embedded nanoreactors and the substrates carrying the brown cargos. Subsequent quencher displacement generates a fluorescent readout. (**B**) In a droplet interface bilayer (DIB) system, mixed substrates in the middle droplet are sorted into the side droplets by their specific reactions with the corresponding nanoreactor variants expressed in the target droplets. We envisage that it is possible to modulate the reaction selectivity in cases where the substrate is threaded from the trans side of the nanoreactor. Different downstream signalling cascades are subsequently triggered in the target droplets. For example, self-assembly of the cargos can be initiated by favourable conditions inside in the left droplet. Alternatively, folded cargos can bind to their target molecules. Also, unmasked reactive groups (e.g. 1,4-aminothiol) can participate in chemical signalling in the right droplet.

2.9 Supplementary materials and methods

2.9.1 Synthesis and characterization of macromolecular substrates

2.9.1.1 Chemicals

Acetic anhydride (Ac_2O), acetonitrile (HPLC grade), 5-azidopentanoic acid, N,N-diisopropylethylamine (DIPEA), diethyl ether, dimethylformamide (DMF), hexafluoro-2-propanol (HFIP), N,N-hydroxybenzotriazole (HOBt), 1-methyl-2-pyrrolidinone (NMP), Nα-Fmoc-Nε-biotinyl-L-lysine, piperidine, triethylamine (TEA), and trifluoroacetic acid (TFA) were purchased from Sigma-Aldrich. 2-(1H-Benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU), Fmoc-Phe-OH, and Rink Amide MBHA resin LL (100-200 mesh) were from Novabiochem. Fmoc-NH-PEG(12)-COOH and Fmoc-Gly-Gly-OH were purchased from Iris Biotech GmbH.

2.9.1.2 Peptide carrier preparation using solid-phase peptide synthesis

The peptide carriers for POC1 and POC2 (Fig. 2.12A and Fig. 2.13A) were synthesized by using manual solid-phase peptide synthesis, adapted from previously reported protocols (*67*). Peptide carriers for POC3 and POC4 were purchased from Peptide Protein Research Ltd.

Peptide elongation was carried out on 0.25 mmol Rink Amide MBHA resin LL (0.37 mmol g^{-1}) by the Fmoc method with HBTU and DIPEA as coupling reagents. The resin was washed and swollen overnight in DMF and Fmoc was removed by treatment with 20% (v/v) piperidine in NMP at room temperature. The carriers were assembled through standard chain elongation by using the reactants Nα-Fmoc-Nε-biotinyl-L-lysine, Fmoc-Phe-OH, Fmoc-NH-PEG(12)-COOH, Fmoc-Gly-Gly-OH (for POC2 peptide carrier) and 5-azidopentanoic acid. For each of the coupling reactions, a solution of 1.0

mmol reactant, 0.95 mmol HBTU, and 2.0 mmol DIPEA in 1.9 mL DMF was added to the resin, which was then shaken for 12 min. After each coupling, the resin was washed with NMP and capped with 1.5 mmol Ac_2O, 0.044 mmol HOBt, and 0.39 mmol DIPEA in 3 mL DMF for 2 min. The capped resin was then washed with NMP again before the next coupling. After the completion of chain assembly, the resin was washed sequentially with DMF and ethanol and dried *in vacuo*. Peptides were cleaved from the resin by treatment with TFA for 2 h. The TFA was evaporated in a stream of nitrogen and the peptide was precipitated with cold diethyl ether, followed by centrifugation, and several trituration treatments with diethyl ether.

2.9.1.3 Peptide carrier purification and characterization

The crude peptide carrier for POC1 or POC2 was dissolved in DMF and purified by preparative RP-HPLC (Dionex UltiMate 3000; Vydac C18 column: 250 × 22 mm, 10-15 μm (particle size); linear gradient: 5-95% eluant B in eluant A over 45 min, flow rate: 15 mL min^{-1}; eluant A: 0.1% TFA in water; eluant B: 0.1% TFA in acetonitrile). The identity and purity of the fractions were determined by analytical LC-MS (Waters LCT accurate-mass time-of-flight instrument (ESI-positive); Chromolith RP-18e column: 50 mm × 2 mm; linear gradient: 5-100% eluant B in eluant A over 8 min, flow rate: 1 mL min^{-1}; eluant A: 0.1% formic acid in water; eluant B: 0.1% formic acid in acetonitrile). The fractions with the correct mass were combined, lyophilized, and characterized by LC-MS and analytical RP-HPLC (Agilent 1260 Infinity HPLC; Polaris C18 column: 150 × 4.6 mm, 5 μm (particle size); linear gradient: 5-95% eluant B in eluant A over 10 min, flow rate: 1 mL min^{-1}; eluant A: 0.1% TFA in water; eluant B: 0.1% TFA in acetonitrile) (Fig. 2.12, B and C, Fig. 2.13, B and C).

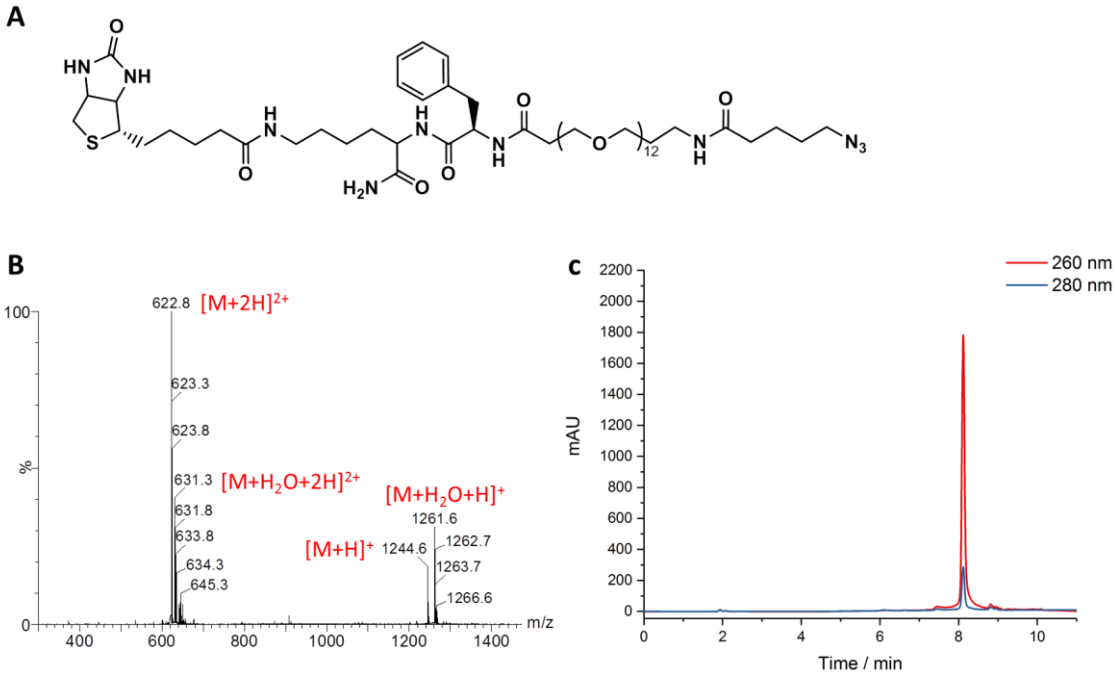


Fig. 2.12. LC-MS and HPLC characterization of the POC1 peptide carrier. (**A**) Structure of the carrier. (**B**) The mass of carrier ($[M+H]^+$) was calculated to be 1243.7 g mol^{-1} and found to be 1244.6 g mol^{-1}. (C) The purity of carrier was confirmed by analytical HPLC.

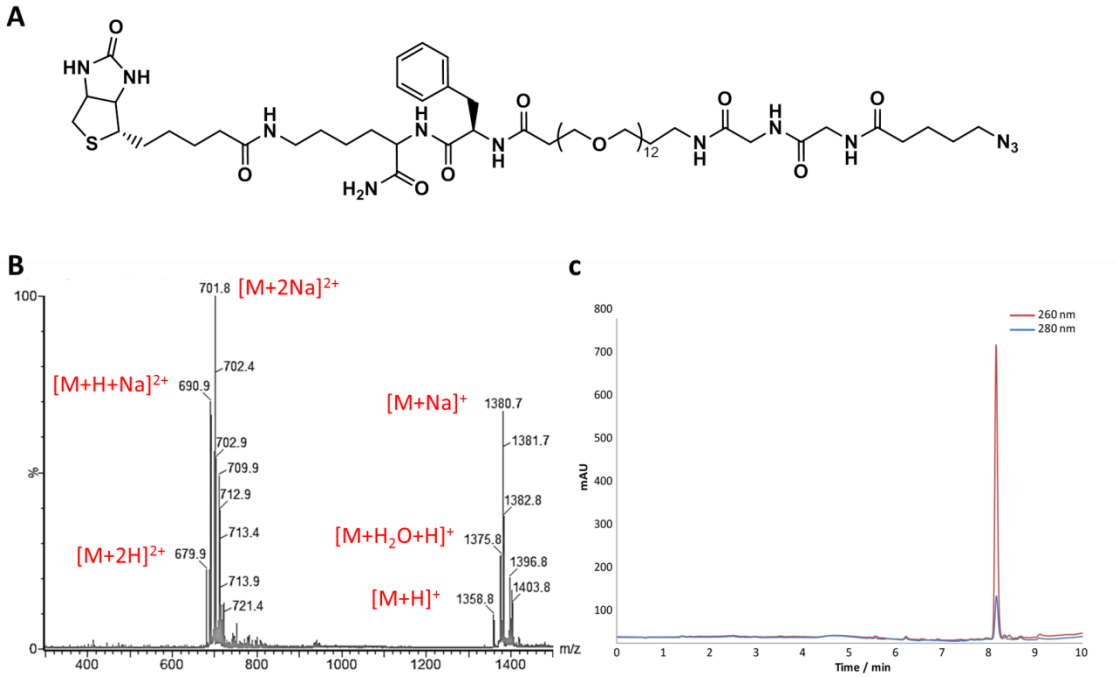
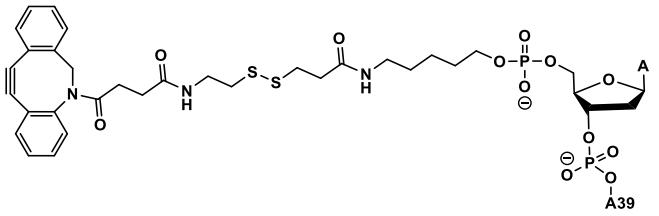


Fig. 2.13. LC-MS and HPLC characterization of the POC2 peptide carrier. (**A**) Structure of the carrier. (**B**) The mass of carrier ($[M+H]^+$) was calculated to be 1358.8 g mol^{-1} and found to be 1358.8 g mol^{-1}. (C) The purity of carrier was confirmed by analytical HPLC.

2.9.1.4 Generation and characterization of macromolecular substrates

Oligonucleotides modified with dibenzocyclooctyne-disulfide (DBCO-SS) at the 5' terminus were purchased from biomers.net.



5' DBCO-SS modification

5' DBCO-SS-A40 oligo (10 µL, 1 mM in MilliQ water) was added to each peptide carrier (1 µL, 10 mM in MilliQ water). The mixtures were left at room temperature for 30 min to form the peptide-oligo conjugates (POC1-4) by a copper-free click reaction (Fig. 2.14).

The conjugates were purified by semi-preparative HPLC (Agilent 1260 Infinity; Supelco Discovery BIO Wide Pore C18 column: 250 × 10mm, 10 µm; linear gradient: 10-90% eluant B in eluant A over 30 min, flow rate: 4.5 mL/min; eluant A: 0.1% TFA in water; eluant B: 0.1% TFA in acetonitrile) and identified by LC-MS (UPLC-MS Waters XEVO G2-QTOF (ESI-negative); ACQUITY UPLC Oligonucleotide BEH C18 Column: 2.1 mm X 50 mm, 1.7 µm (particle size); linear gradient: 0-70% eluant B in eluant A over 8 min, flow rate: 0.2 mL min^{-1}; eluant A: 8.6 mM TEA, 200 mM HFIP in 5% methanol/water (v/v); eluant B: 20% eluant A in methanol).

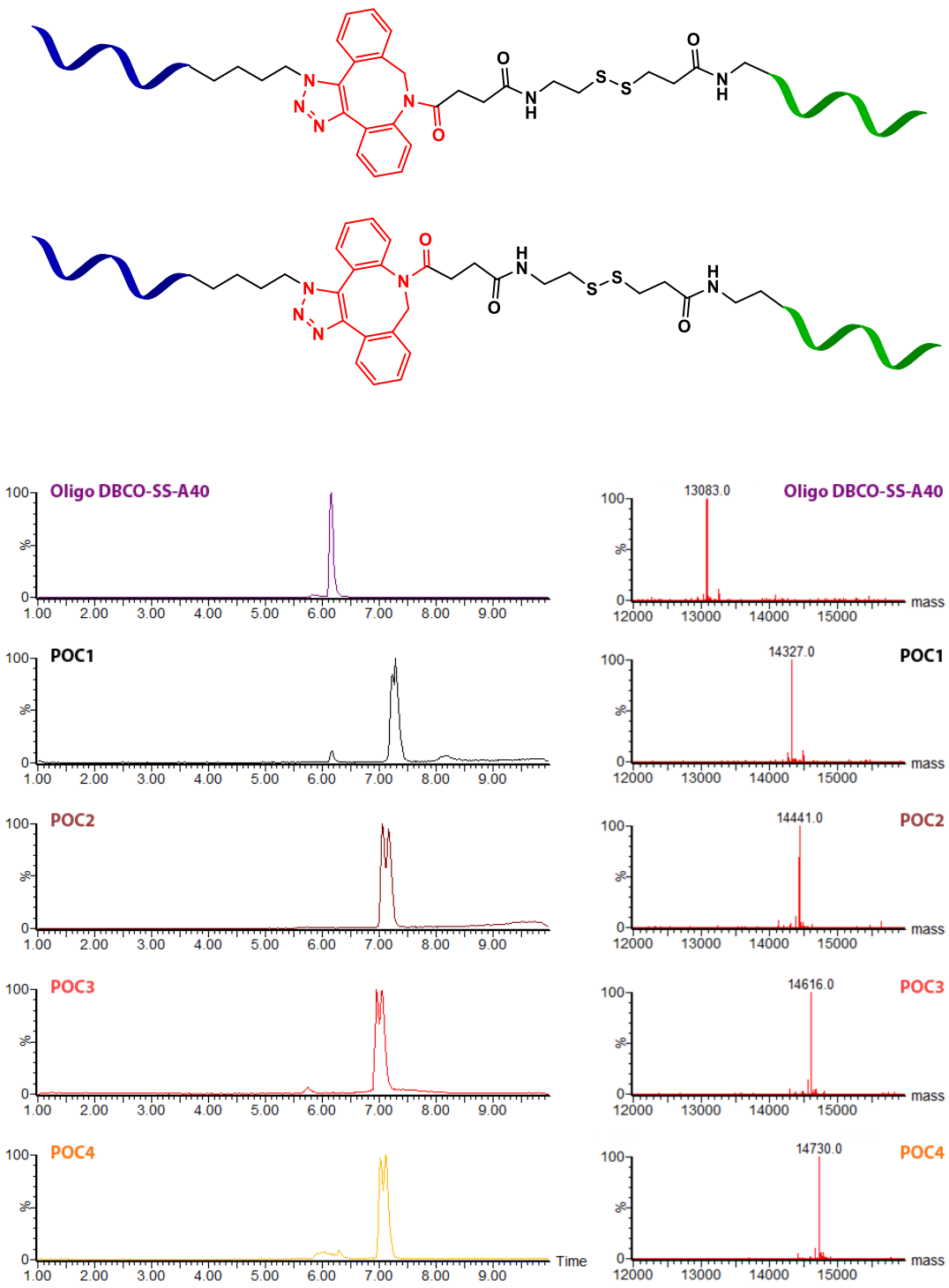


Fig. 2.14. LC-MS characterization of macromolecular substrates POC1-4. The two peaks in each chromatogram have the same mass, corresponding to the two regioisomeric conjugates formed by the copper-free click chemistry. The mass of POC1 was calculated to be 14327 g mol^{-1} and found to be 14327 g mol^{-1}. The mass of POC2 was calculated to be 14440 g mol^{-1} and found to be 14441 g mol^{-1}. The mass of POC3 was calculated to be 14615 g mol^{-1} and found to be 14616 g. The mass of POC4 was calculated to be 14729 g mol^{-1} and found to be 14730 g mol^{-1}.

2.9.1.5 Preparation of cysteine-containing αHL monomers

The construction of single-cysteine nanoreactors 113, 115, 117, 119, 121 and the double-cysteine nanoreactor 115/117 based on pT7-αHL-D8H6, which encodes the wild-type αHL with an octa-aspartate tail and a His-tag at the C terminus, have been previously reported (*91*).

Plasmid pT7-αHL-143C-D8H6 encodes an αHL mutant containing a cysteine at position 143 as well as an octa-aspartate tail and a His-tag at the C terminus. It was generated from pT7-αHL-D8H6 by *in vivo* homologous recombination. Two sets of PCR reactions were carried out. For the first reaction, the template was first linearized by NdeI (New England Biolabs) before PCR with the forward mutagenic primer: 5'-CAAATGTTTCGATTTGTCATACACTGAAATATGTTC-3' and the reverse nonmutagenic primer (SC47): 5′-CAGAAGTGGTCCTGCAACTTTAT-3′. For the second reaction, the template was linearised by HindIII (New England Biolabs) before PCR with the reverse mutagenic primer: 5'-GAACATATTTCAGTGTATGACAAATCGAAACATTTG-3' and the forward nonmutagenic primer (SC46): 5′-ATAAAGTTGCAGGACCACTTCTG-3′. The linearized templates and primers were mixed with 1XPhusion Flash HF Mastermix (New England Biolabs) and put through the following cycling program: 94°C for 5 min, then 30 cycles of 94°C (30 s), 50°C (30 s), 72°C (30 s), and then 50°C for 5 min. After PCR, 5 μL of each reaction were mixed and transformed into E. coli XL10-Gold cells (Agilent). The transformed cells were allowed to grow at 37°C overnight on LB (Luria Broth)-carbenicillin plates. Plasmid DNA was isolated from colonies by using the QIAprep Spin Miniprep Kit (QIAGEN). Successful mutagenesis was confirmed by DNA sequencing.

Similarly, pT7-αHL-115C143C-D8H6 was made from pT7-αHL-115C-D8H6 with the same two mutagenic primers. pT7-αHL-123C-D8H6 was made from pT7-αHL-WT-D8H6 with the two mutagenic primers: 5'-CGGATTCAACGGTTGTGTTACTGGTGATGATACAGG-3' (Forward); 5'-CCTGTATCATCACCAGTAACACAACCGTTGAATCCG-3' (Reverse).

2.9.1.6 Preparation of αHL heptamers containing cysteine(s) on one of the seven subunits

Engineered αHL polypeptides were expressed by using a commercial *in vitro* transcription-translation (IVTT) kit: *E. coli* T7 S30 Extract System for Circular DNA (Promega). To suppress transcription by *E. coli* RNA polymerase, the T7 S30 extract provided in the kit was treated with rifampicin prior to use (1 μg mL^{-1}, final concentration). A standard reaction comprised: DNA template (3.2 μg), amino acid mix minus methionine (supplied with the kit, 5 μL), S30 premix without amino acids (supplied with the kit, 20 μL), [^{35}S]methionine (2 μL, 1,200 Ci $mmol^{-1}$, 15 mCi mL^{-1}, MP Biomedicals), T7 S30 extract (supplied with the kit. 15 μL), and nuclease-free water to a final volume of 50 μL. To make heteroheptamers, a mixture of plasmids encoding the WT αHL and the mutant monomer were mixed in a ratio 6:1 (WT: mutant). The IVTT mixture was incubated at 37°C for 1 h.

Heptamerization was carried out by the addition of rabbit erythrocyte membranes (3 μL, ~1 mg protein mL^{-1}) to the IVTT reaction mixture (50 μL), followed by incubation at 37°C for another 1 h. The mixture was then centrifuged for 10 min at 25,000 × g. The supernatant was removed and the pellet resuspended in MBSA buffer (200 μL, 10 mM 3-morpholinopropane-1-sulfonic acid (MOPS), 150 mM NaCl, 1 mg mL^{-1} bovine serum albumin, pH 7.4). The wash with MBSA was repeated before the

pellet was resuspended in 2X Laemmli sample buffer (50 μL) and electrophoresed in a 5% SDS polyacrylamide gel at 70 V for 15 h.

The αHL heteroheptamers containing different numbers of mutant subunits were separated in the gel based on their different electrophoretic mobilities, which were determined by the number of octa-aspartate (D8) tails (*92*). The top and bottom bands corresponded to WT_7 and $(mutant\text{-}D8H6)_7$, respectively. The second band from the top was the desired heteroheptamer containing a single mutant subunit.

To extract heptameric pores, the gel was first dried without fixation on Whatman 3M filter paper under vacuum for 5 h at room temperature. After visualization by autoradiography with Kodak BioMax MR film, the desired bands were cut from the gel with a scalpel. Each excised band was rehydrated in TE buffer (300 μL, 10 mM Tris·HCl, 1 mM ethylenediaminetetraacetic acid (EDTA), pH 8.0) for 1 h at room temperature. The filter paper was then removed, and the rehydrated gel was macerated with a pestle. The resulting slurry was filtered through a 0.2 μm hydrophilic membrane filter (Proteus Mini Clarification Spin Column, Generon). The filtrate was stored in 10 μL aliquots at -80°C.

2.9.2 Single Channel Recordings

2.9.2.1 General

1,2-Diphytanoyl-sn-glycerol-3-phosphocholine (DPhPC) was purchased from Avanti Polar Lipids. Unless otherwise stated, all other chemicals were purchased from Sigma-Aldrich.

Planar bilayer recordings were performed following the method established by Montal and Mueller(*93*). Two Delrin compartments were separated by a 25 μm-thick Teflon film containing an aperture (60 μm in diameter). The aperture was pre-treated with 1% (v/v) hexadecane in pentane. Each compartment was then filled with buffer

(500 μL, 2 M KCl, 20 mM N-(2-hydroxyethyl)piperazine-N'-(4-butanesulfonic acid) (HEPBS), 20 μM EDTA, pH 8.5). The trans compartment contained an additional 5 mM DTT. Drops of DPhPC in pentane (5 mg mL^{-1}) were added to both compartments. Repetitive up-and-down pipetting of the buffer solution resulted in the formation of a lipid bilayer across the aperture. A transbilayer potential was applied with two Ag/AgCl electrodes, each contained within a salt bridge formed from 3 M KCl in 3% (w/v) low-melt agarose.

Ionic currents were recorded by using a patch clamp amplifier (Axopatch 200B, Axon Instruments) with a 4-pole low-pass Bessel filter (80 dB/decade) at room temperature (20 ± 1°C). Signals were digitized with a Digidata 1320A digitizer (Molecular Devices), connected to a computer running the pCLAMP 10.3 software suite (Molecular Devices). Unless stated otherwise, the signal was filtered with a corner frequency of 5 kHz and sampled at 25 kHz.

2.9.2.2Monitoring turnovers

The substrate (400 μM in 1 μL MilliQ water) was added to a solution of traptavidin (Kerafast) (40 μM in 10 μL phosphate-buffered saline, pH 7.4) and incubated at room temperature (20 ± 1°C) for 15 min to form the traptavidin-tagged substrate. The cis compartment of a planar bilayer recording apparatus containing the αHL nanoreactor was stirred until a single pore inserted into the bilayer. The traptavidin-tagged substrate (3 μL, 36 μM) was then added to the cis compartment, which contained 500 μL of recording buffer (2 M KCl, 20 mM HEPBS, 20 μM EDTA, pH 8.5). The trans compartment contained an additional 5 mM DTT. To drive the macromolecular substrate into the nanopore, a transmembrane potential of +150 mV was applied to the trans side. The threaded substrate reduced the current to an $I_{res\%}$ of ~33% (see Table 2.2 for each substrate-nanoreactor combination). After the thiol-

disulfide interchange between the substrate and the nanoreactor, the current increased as one segment of the substrate departed the nanoreactor ($I_{res\%}$ of 39%-53% for αHL-oligo adducts and ~24-56% for αHL-peptide adducts, Table 2.2). DTT later released the adduct and returned the nanoreactor to its initial state ready for another round of thiol-disulfide interchange.

2.9.2.3 Single-channel data analysis

Current traces were idealized by using Clampfit 10.3 (Molecular Devices).

The idealized data were analyzed with QuB 2.0 software (www.qub.buffalo.edu) (*94*). Dwell time analysis and rate constant estimations were performed by using the maximum interval likelihood (MIL) algorithm of QuB (*76*).

Chapter 3 Molecular Hopper

3.1 Introduction

Processivity lies at the heart of biological machines. A replicative DNA polymerase can incorporate thousands of nucleotides before dissociating from its template (*95*). Molecular motors, such as kinesin and dynein, travel directionally along microtubules over hundreds of steps without detaching from the track (*96–98*). For years, scientists have been trying to build moving molecules that resemble their biomolecular counterparts but use simpler components (*99*). The ultimate goals are to achieve true processivity, which can be defined as directional motion without leaving a track and the performance of useful work such as the transport of a cargo. Ideally, a synthetic system should exhibit the reversibility of stepping seen in various biological systems (*100, 101*) to enable the direction of motion to be switched through external control.

We report the design of a one-legged molecular hopper that is ratcheted by dynamic covalent chemistry along a protein track (Fig. 3.1) with robust processivity (Table 3.1). Further, the direction in which the hopper moves is subject to external control by an electrical potential. To highlight the advances presented by our molecular hopper, a selection of its properties was compared to the best achievements across various reported examples from three different classes.

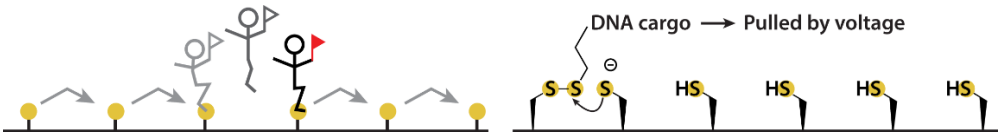


Fig. 3.1. Schematic representation of the molecular hopper. A hopper carrying a cargo (red flag) moves along a track by means of consecutive thiol-disulfide interchange reactions. The overall direction is set by the applied potential.

Table 3.1. Comparison of the hopper with three classes of 'moving molecules' [a].

	Small molecule "walkers"		Macromolecular walkers	
	HOPPER (This work)	**Previous small-molecule walkers** (*91*, *102–108*)	**Previous DNA-walkers** (*109–116*)	**Biomolecular walkers** (*95–98*, *100*, *117*)
Synthesis	Straightforward	Multistep chemical syntheses	DNA synthesis and assembly	Biological
Track chemistry	Simple, transferable between walkers	Complex for directional motion	Moderately simple, not transferable	Complex, not transferable
Step size	0.7 nm	< 1 nm	Typically ~10 nm (5 nm (*116*))	Typically ~10 nm (Kinesin ~ 8 nm)
Processivity[b]	Processive (100s of steps[c])	Limited (<10 steps demonstrated[d])	Processive (<100 steps demonstrated)	Processive (100s of steps)
Directionality[b]	Directional[e]	Limited directionality	Directional (burnt-bridge mechanism)	Directional
Rate	Seconds per step[f]	Minutes to days per step	10s of minutes per step	Milliseconds per step
Real time step detection	Yes	No	In some cases	Yes
Autonomy[d]	Yes	Generally no (Autonomous (*106*))	May require fuel	Requires ATP
Reversibility (ability to make 180° turns)	Fully reversible	Not reversible	Not reversible	Limited reversibility
External control	Yes	Limited	Limited	No
Track length	Limited based on protein nanopore (nm scale)	Limited by chemical synthesis (nm scale)	Limited by DNA biotechnology (μm scale)	μm to mm scale
Attachment and detachment points	Both precisely defined	Pre-attachment	Precisely defined	Biologically defined
Ability to carry molecular cargos	Yes[g]	Limited	Yes[g]	Yes

[a] Here, we only compare molecular 'walkers'. Additional 'moving molecules' include: sleds (*118*), rotors (*119, 120*), pumps (*121*).

[b] Processivity is defined as directional stepping without leaving the track. A directional step has a faster forward stepping rate than the backward rate, i.e. $K = k_{forward}/k_{backward} > 1$.

In previous small-molecule systems, detachment from the track has been prevented through various strategies such as the maintenance of covalent linkages between walkers and their tracks (*102*, *105*, *106*) and the exploitation of mechanically interlocked architectures (*103*, *104*). For DNA-based walker systems, programmable hydrogen-bonding allows the walkers to remain hybridized with track components (*114*).

Directionality has been demonstrated in small-molecule systems for short excursions on tracks terminating with thermodynamic sinks (*91*, *105*, *106*), or on tracks with switchable conformations (*107*), or through alternating orthogonal walker-track chemistries (*102*, *108*), or through sequential installation and removal of stoppers on a track (*103*). Unidirectional motion of DNA-based walkers has been achieved through the 'burnt-bridge' mechanism by cycles of DNA hybridization and hydrolysis at the cost of reversibility (*110*, *111*, *114*).

[c] Up to 5 steps in each direction were demonstrated due to limited track length. 100s of processive steps were achieved through 'flipping the potential' after the hopper reached the final foothold.

[d] Autonomy means motion without chemical fuel or facilitator molecules (*91*).

Autonomous eight-step walking has been demonstrated with a nine-foothold track terminating with a thermodynamic sink through dynamic covalent chemistry—reversible Michael addition of secondary amines to an α-methylene-4-nitrostyrene unit (*106*). The fastest forward walking rate was reported to be 5×10^{-5} s^{-1} (~5.6 h per step). The equilibrium constants (K) for the first six steps were reported to be 1, in line with the initial random walking on the track. The last two steps have K values of 1.3 and 2.5, respectively, to which the overall biased directionality was attributed. In comparison, the present hopper has a hopping rate of ~22 s per step and a forward equilibrium constant (K) of ~5 at +150 mV and ~20 at -150 mV.

[e] All steps demonstrated with the hopper have an equilibrium constant larger than one, and hence are directional. On the five-cysteine track at -150 mV, hopping from 117 to 115 has a $K = k_{117\text{-}115}/k_{115\text{-}117} = 22$, which indicates that the observed back-stepping was due to a slow stepping rate for the next step to the downstream foothold 113.

[f] On the five-cysteine track, the mean step time is ~22 s at ±150 mV.

[g] Constrained by the current nanopore-based track, macromolecular cargos movable by the hopper system are limited to linear molecules such as oligonucleotides and polypeptides. DNA-walkers moving on a surface, on the other hand, are capable of transporting cargos such as nanoparticles (*112, 115*).

3.2 Directional control of a processive molecular hopper

The track is built inside a protein nanopore, α-hemolysin (αHL), and consists of a series of cysteine footholds facing the lumen of the transmembrane β barrel (Fig. 3.2A). The cysteines are evenly spaced along a β strand with an average inter-foothold distance of 6.8 Å (Cα-Cα) and an average vertical spacing of 5.6 Å.

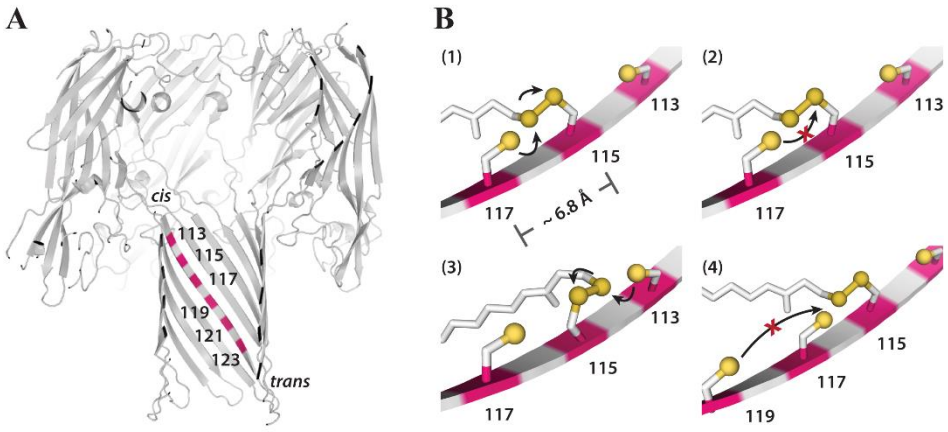


Fig. 3.2. A molecular hopper on a protein track. (**A**) A six-foothold track comprising odd-numbered cysteine residues on a β strand inside the αHL protein nanopore. (**B**) The applied potential exerts a force on the DNA cargo, which helps to align the three sulfur atoms (yellow) participating in the interchange. The collinear geometry promotes hopping (1) (movie S1) but not the formation of an intrastrand disulfide, which would release the hopper from the track (2). Occasionally, backstepping is observed (3). Overstepping (4) does not occur.

The hopper moves in the direction in which the DNA cargo has been oriented by an applied potential, by employing consecutive thiol-disulfide interchange reactions. To execute the S_N2 reaction, the three participating sulfur atoms must align in a near linear configuration (*25*, *26*, *49*). Under the applied potential, the DNA inside the barrel is pulled in the electric field with a force of ~10 pN (see section 3.4). The force sets the overall direction of motion by flipping the DNA (see below) and helps orient the disulfide for cleavage by the neighbouring downstream cysteine thiolate, which moves the hopper one step forwards, although other forces can contribute to the forward

motion (Fig. 3.2B). Back-stepping is disfavoured and over-stepping is impossible. Release of the hopper from the linear track was not observed, presumably because the track is too rigid to accommodate the resultant disulfide bridge between adjacent footholds on the same β strand. In short, each step is chemically directional as the hopper's 'foot' is positioned to favour the forward reaction. Further, the motion is autonomous, requiring no chemical fuel.

Under +150 mV, the hopper was delivered to the track from the cis compartment as a hopper-carrier conjugate (Fig. 3.3A) capped with a single traptavidin, which was arrested at the pore entrance (Fig. 3.3B). The disulfide in the construct reacted regioselectively with Cys-115 releasing the carrier and placing the hopper-DNA cargo on the starting foothold (Fig. 3.3B and Fig. 3.4).

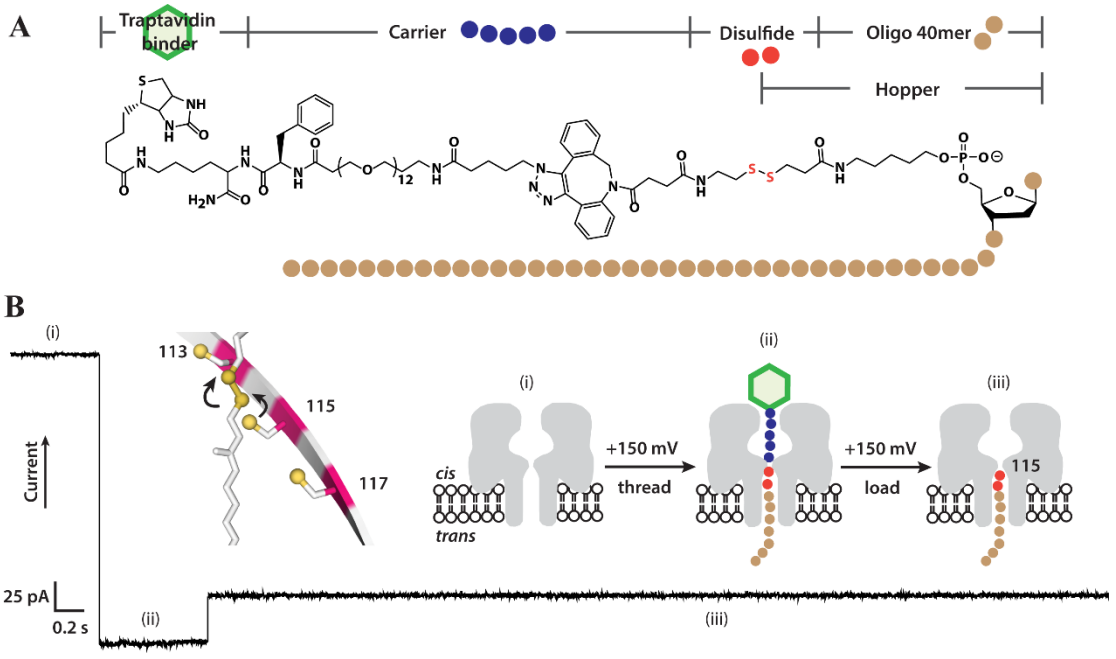


Fig. 3.3. Loading the molecular hopper on a protein track. (**A**) The hopper enters the nanopore as a carrier-hopper disulfide conjugate. (**B**) Under +150 mV, a hopper-carrier conjugate capped with traptavidin was pulled from the cis compartment into an αHL nanopore containing cysteines at positions 113, 115, 117, 119, and 121 in one of the seven subunits. The resultant blockade reduced the ionic current from (i) to (ii). Reaction of the disulfide in the hopper-carrier with Cys-115 covalently attached the hopper to the track, and the ionic current increased to (iii).

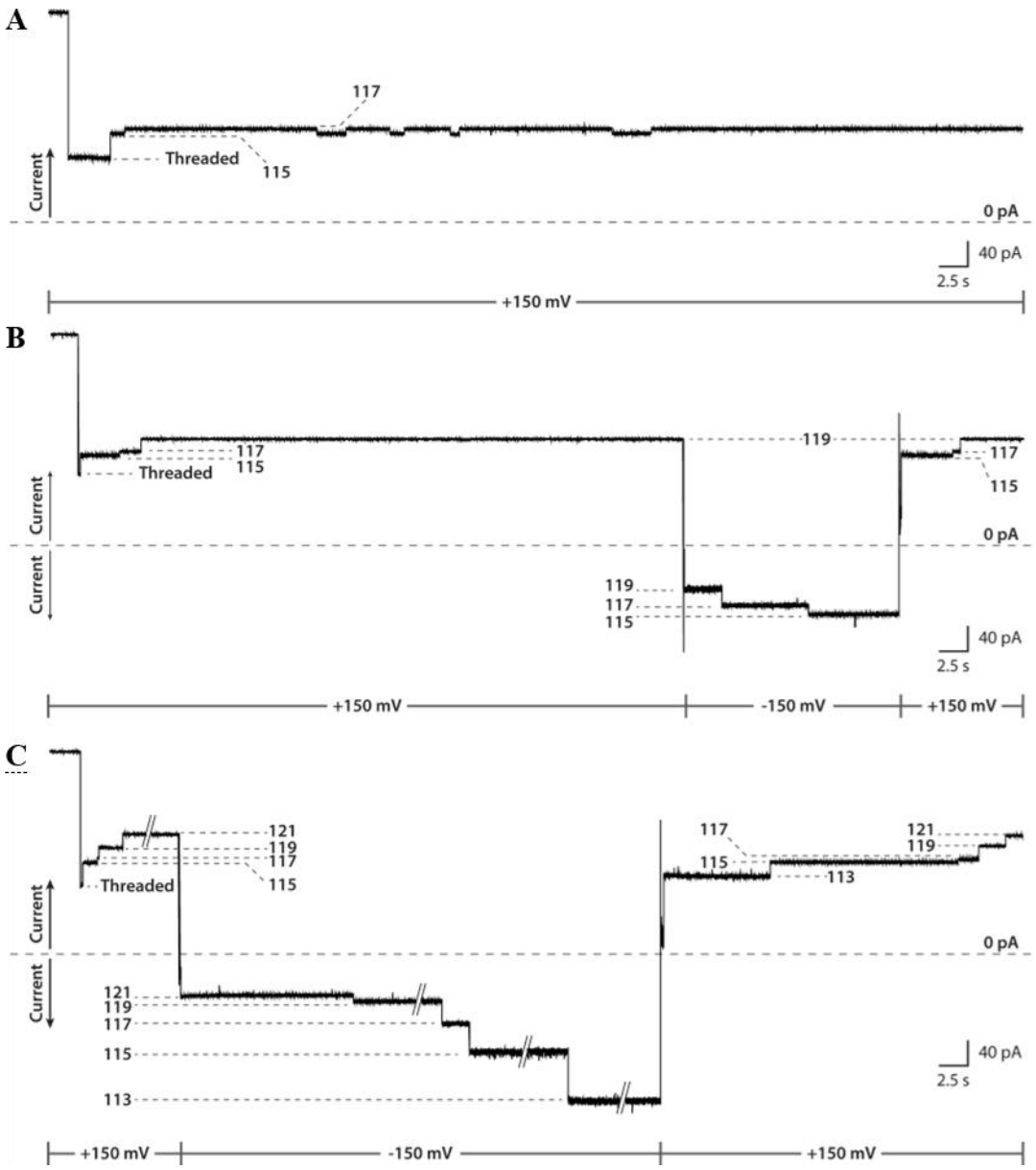


Fig. 3.4. Loading hopper 1 onto different tracks. Under +150 mV, a hopper-1-carrier conjugate capped with a traptavidin threaded from the cis side of the nanopore before reacting with Cys-115, which covalently attached the hopper to (**A**) a two-cysteine track (Footholds: Cys-115, 117); (**B**) a three-cysteine track (Footholds: Cys-115, 117, 119); (**C**) a five-cysteine track (Footholds: Cys-113, 115, 117, 119, 121). The movement of the hopper was tracked from locations ascertained from the residual current passing through the nanopore. With the five-cysteine track, only three steps were seen in the first excursion at +150 mV. Following the four-step hopping at -150 mV, the second excursion at +150 mV on the same track exhibited four steps, indicating that the initial attachment was on Cys-115. The trace was filtered at 200 Hz. Conditions: 2 M KCl, 20 mM HEPBS, 20 μM EDTA, pH 8.5, 20 ± 1 °C.

The location of the hopper was ascertained from the residual current passing through the nanopore, which reflected the length of the DNA located within the β barrel when the hopper was at a particular foothold (Fig. 3.4). By monitoring current changes, we followed the stepwise hopping motion at the single-molecule level in real-time.

The voltage-controlled hopping motion was directional and processive. On a track containing five cysteine footholds (113, 115, 117, 119, 121), a hopper carrying an oligoadenosine 40-mer (A40, hopper 1) moved cis to trans under +150 mV, and trans to cis under -150 mV (Fig. 3.5 and Fig. 3.6). When the hopper reached a terminal foothold, the sign of the applied potential was reversed in order to reorient (flip) the DNA cargo, and hence the hopper. Alternation between positive and negative potentials repeatedly drove the hopper towards the trans or the cis end of the track.

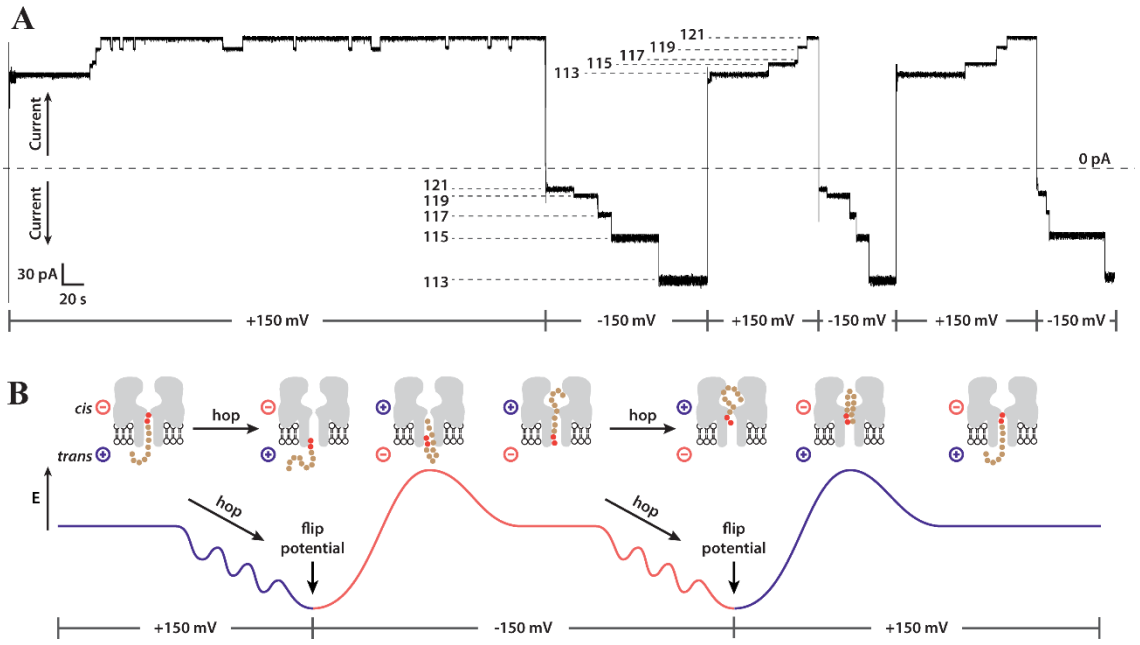


Fig. 3.5. Monitoring individual hopper steps. (**A**) With a five-cysteine track, four hopping steps were observed at ±150 mV. Every forward step moved part of the DNA cargo outside the β barrel, producing an increase in conductance. Alternation of the applied potential drove the hopper repeatedly up and down the track. (**B**) A hypothetical free energy diagram (not to scale) of the controlled hopping motion.

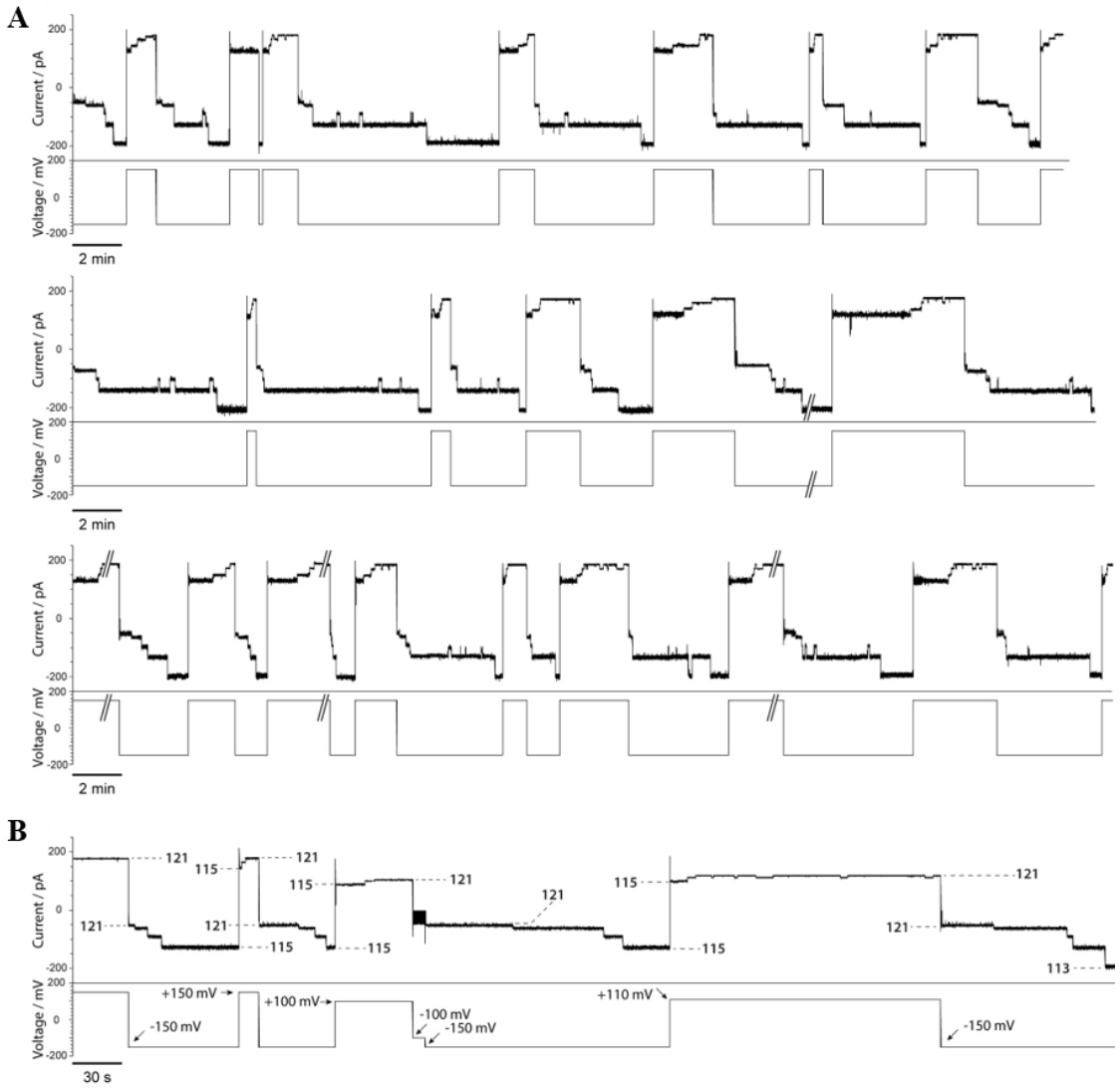


Fig. 3.6. Hopping cycles on the five-cysteine track. After hopper 1 was loaded onto the five-cysteine track (Footholds: Cys-113, 115, 117, 119, 121), it moved under external control for >20 complete cycles (from 113 to 121 at +150 mV and then from 121 to 113 at -150 mV) during a single recording. (**A**) Complete cycles at ±150 mV are shown, from which the kinetic rate constants were derived by using QuB (see section 3.3). (**B**) External manipulation of voltage to various values (e.g. +100 mV) was applied at non-terminating footholds as well (e.g. Cys-115) in the same experiment. The total number of steps was 282 (only back-steps at the non-terminating footholds were counted; steps after reaching the final footholds and before the voltage switch were discounted). Traces were filtered at 200 Hz. Conditions: 2 M KCl, 20 mM HEPBS, 20 μM EDTA, pH 8.5, 20 ± 1 °C.

For the DNA to experience a force, at least one negatively-charged phosphodiester bond must lie within the electric field, which drops along the length of the pore's β barrel (*56*). Therefore, in the present nanopore construct, the length of the track is limited and voltage-controlled hopping was demonstrated with up to six footholds (Fig. 3.7). The hopping direction could be changed by reversing the applied potential at any foothold, showcasing the complete control over directionality and the ability to move a hopper back to the initial foothold after an outing.

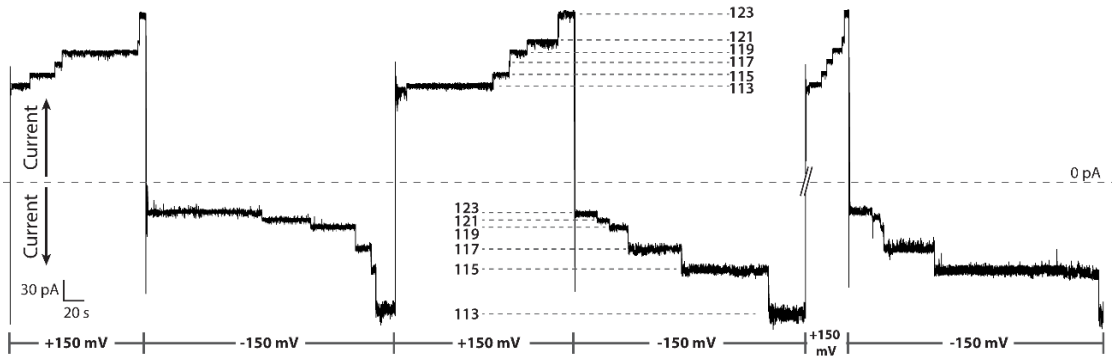


Fig. 3.7. Hopping on a six-cysteine track. Five hopping steps were recorded for hopper 1 on a six-cysteine track (Footholds: Cys-113, 115, 117, 119, 121, 123) at ±150 mV. The trace was filtered at 200 Hz. Conditions: 2 M KCl, 20 mM HEPBS, 20 μM EDTA, pH 8.5, 20 ± 1 °C.

In the present system, the applied potential provides an external energy source to produce directional motion (Fig. 3.5B, see section 3.4). Limited by bilayer stability, the longest records of processive hopping were documented with hopper 1, which completed 249 forward steps in 93 min on a five-cysteine track (113-121) with a mean dwell-time of ~22 s per foothold (Fig. 3.6A). Dissociation of the hopper from the track was never seen (n >30 outings on different tracks), which implies that substantial improvements on step numbers would be achieved if the stability of the bilayer were improved. In comparison, previous synthetic small-molecule walkers moved directionally for less than 10 steps (*91*, *99*). Wild-type kinesins typically exhibit a mean step number of 75–175 before dissociation (*96*, *97*).

3.3 Hopping kinetics

The hopping rates for each of the four steps on the five-cysteine track were derived for both the cis-to-trans and trans-to-cis directions at pH 8.5 and displayed differences of less than 40-fold (0.0081 to 0.30 s^{-1}, Table 3.2, Table 3.3, Table 3.4). Because a thiolate is the reactive nucleophile in disulfide interchanges, the rate differences might arise from variations in the pK_a values of the foothold thiols, which will be affected by neighbouring residues. Previously, an arsenic(III) walker showed an up to a 50-fold difference in attachment rates with the footholds on the same five-cysteine track (113-121) at pH 8.0 (*91*). In the future, tracks of thiols might be engineered with optimized inter-foothold distances and enhanced chemical reactivity to speed up the hopping process. Alternatively, the properties of the reactive sulfur atom in the hopper might be manipulated by flanking functional groups. With both the two-cysteine and the three-cysteine tracks, the influence of voltage on hopping was examined at ±100 mV, ±150 mV, and ±180 mV (Table 3.5, Table 3.6, Table 3.7). The rates showed weak non-exponential voltage-dependences suggesting that the applied potential might not be the only source of propulsion (see section 3.4). However, an electrical potential is essential 1) to flip the DNA over a large barrier to set the direction of motion; 2) to aid in the orientation of the three participating sulfur atoms to favor "forward" reactions over back-steps. With respect to the latter, the estimated "effective concentrations" of participating downstream thiols are not especially high (see section 3.3.4) (*122*) and indeed need not be to produce overall forward motion (see below).

Although disfavoured, back-stepping was occasionally detected and attributed to conformational lability of the hopper within the nanopore even under an applied potential. During a recording with hopper 1 on a five-cysteine track (113-121), there were 33 backward steps on the non-terminating footholds out of 282 steps in total

(12%). Of the 33 backward steps, 29 occurred from 115 to 117 at -150 mV (Table 3.3). Despite the large forward equilibrium constant for 117 to 115 ($K = k_{117\text{-}115}/k_{115\text{-}117} = 22$), 115-117 back-stepping is observed, because of the comparatively slow forward movement to the next foothold Cys-113 (at -150 mV, $k_{115\text{-}117} = 0.0094\ s^{-1}$; $k_{115\text{-}113} = 0.0081\ s^{-1}$, Table 3.2 and Table 3.3). Back-stepping was observed when the hopper was left on a terminal foothold as no forward footholds remained (Fig. 3.5A). These back-steps were quickly reversed by the hopper, which preferentially resided at the final station ($K = k_{119\text{-}121}/k_{121\text{-}119} = 5.2$, Table 3.4). The overall motion of the hopper is governed by the product of the K values for each step. A modest value ($K > 1$) at each step produces a considerable overall tendency towards forward movement.

3.3.1 Hopping rates at ±150 mV for the four steps of the five-cysteine track.

Table 3.2. Hopping rates (k) and mean dwell times ($<\tau>$) derived by QuB[a] for hopper 1 on the five-cysteine track (Footholds: Cys-113, 115, 117, 119, 121) at ±150 mV[b].

Forward step at +150 mV	**113-115**	**115-117**	**117-119**	**119-121**
k / s^{-1}	0.027 ± 0.006	0.084 ± 0.017	0.30 ± 0.06	0.23 ± 0.05
$<\tau>$ / s	37 ± 8	12 ± 2	3.3 ± 0.7	4.3 ± 0.9
n	22	23	21	21
Forward step at -150 mV	**121-119**	**119-117**	**117-115**	**115-113**
k / s^{-1}	0.086 ± 0.017	0.047 ± 0.010	0.21 ± 0.03	0.0081 ± 0.0017
$<\tau>$ / s	12 ± 2	21 ± 5	4.7 ± 0.7	124 ± 26
n	22	21	50	21

[a] Dwell time analysis and rate constant estimations were performed by using the maximum interval likelihood algorithm of QuB. An arbitrary duration of 1 s was given to the final foothold on either direction (Cys-121 for cis to trans hops at +150 mV or Cys-113 for trans to cis hops at -150 mV) to complete the kinetic model (Fig. 3.8).

[b] Only steps from complete cycles (from 113 to 121 at +150 mV and then from 121 to 113 at -150 mV) (Fig. 3.6A) were used to fit the kinetic model (Fig. 3.8). Data were collected from a single αHL pore.

Table 3.3. Hopping rate (*k*) and mean dwell time (<τ>) derived by QuB[a] for the 115-117 backward steps of hopper 1 on the five-cysteine track (Cys-113, 115, 117, 119, 121) at -150 mV[b].

	115-117 (-150 mV)	**119-121 (-150 mV)**	**117-115 (+150 mV)**	**115-113 (+150 mV)**
***k* / s^{-1}**	0.0094 ± 0.0019	/	/	/
<τ> / s	106 ± 21	/	/	/
n	29	1	2	1

[a] Dwell time analysis and rate constant estimations were performed by using the maximum interval likelihood algorithm of QuB. Apart from 115-117 backward steps, which occurred 29 times, the other backward steps were rare and no meaningful rate constants could be derived. QuB was hence set to treat the forward steps as irreversible (Fig. 3.8).

[b] Data were collected from a single αHL pore.

Table 3.4. Hopping rates (*k*) and mean dwell times (<τ>) derived by QuB[a] for hopper 1 on the last two footholds (Cys-119, 121) of the five-cysteine track (Cys-113, 115, 117, 119, 121) at +150 mV[b].

Step at +150 mV	**119-121**	**121-119**
***k* / s^{-1}**	0.23 ± 0.03	0.044 ± 0.005
<τ> / s	4.3 ± 0.5	23 ± 3
n	75	75

[a] Dwell time analysis and rate constant estimations were performed by using the maximum interval likelihood algorithm of QuB. Since there was no 119-117 back-stepping during any hopping cycles, a two-state kinetic model was used to determine the stepping kinetics on the last two footholds (Fig. 3.9).

[b] Data were collected from a single αHL pore.

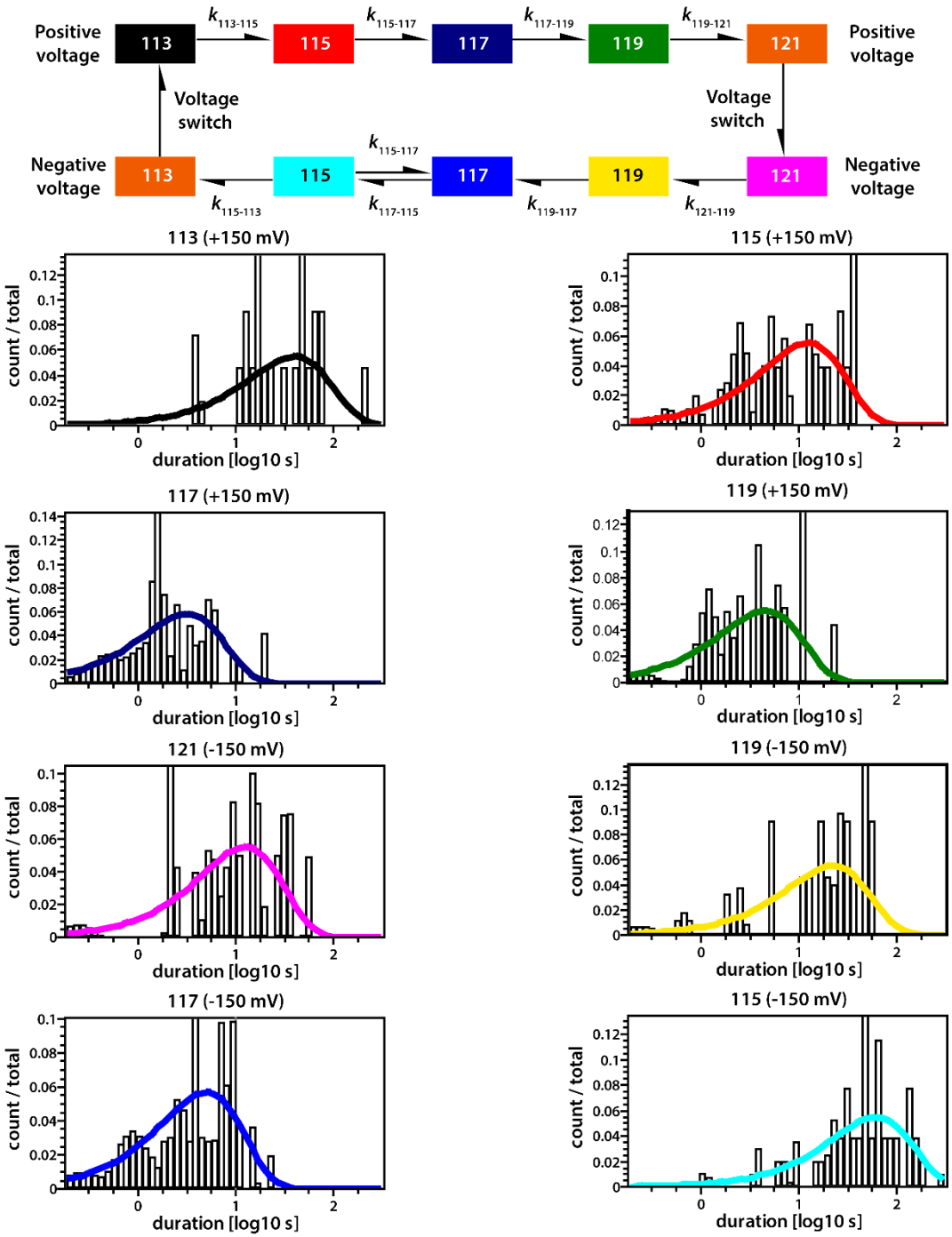


Fig. 3.8. Kinetic model and duration histograms for the four steps on the five-cysteine track (Footholds: Cys-113, 115, 117, 119, 121). The kinetic model (top) links the four steps at positive voltages with the four at negative voltages to form a closed cycle. Two voltage-switch transitions were built into the model by assigning an arbitrary duration of 1 s to the final foothold of the track (Cys-121 at +150 mV; Cys-113 at -150 mV). QuB was set to treat the forward steps as irreversible apart from 117-115, where back-stepping occurred 29 times. Back-steps after reaching the final footholds were discounted. The derived kinetic rate constants are given in Table 3.2 and 3.3. Data were collected from a single αHL pore.

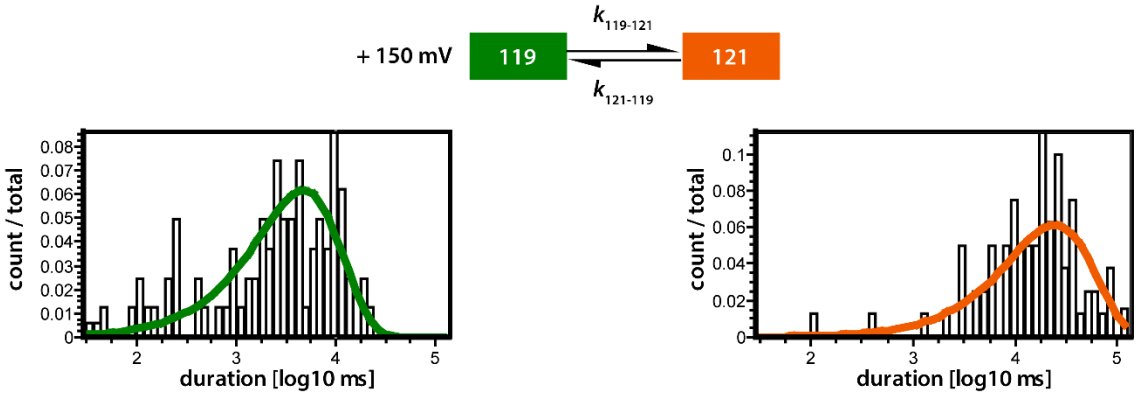


Fig. 3.9. Kinetic model and duration histograms for the forward and backward steps on the last two footholds (Cys-119, 121) of the five-cysteine track (Footholds: Cys-113, 115, 117, 119, 121) +150 mV. The forward and backward rate constants were derived by using the maximum interval likelihood algorithm of QuB, according to a two-state kinetic model (top). The derived kinetic rate constants are given in Table 3.4. Data were collected from a single αHL pore.

3.3.2 Hopping rates at ±100 mV, ±150 mV, ±180 mV for the one step of the two-cysteine track

A two-cysteine track consisting of Cys-115 and Cys-117 was used to study the rates of forward and backward hopping under various applied potentials (±100 mV, ±150 mV, ±180 mV). Hopper 1 moved forward from foothold 115 to 117 under a positive applied potential and from foothold 117 to 115 under a negative applied potential. Occasional back-stepping was also observed. Rate constants (k) were derived for hopping steps under both positive and negative voltages (Table 3.5). The equilibrium constant K was then calculated for each potential, where $K = k_{forward}/k_{backward}$.

The weak voltage dependence of K under the positive applied potentials is consistent with the DNA cargo being driven towards a higher potential. The field drops primarily across the β barrel (*56*) and four negatively charged phosphodiesters are located inside the barrel based on PyMOL modelling (Fig. 3.10B). The trend is less clear at negative potentials, but in this case only one or zero phosphodiesters are located

inside the barrel (Fig. 3.10B). Although there is a weak electric field in the αHL vestibule (*56*), the directionality of hopping under negative applied potentials is more likely to arise from DNA coiling in the cis vestibule which disfavours backward movement of the hopper (see section 9).

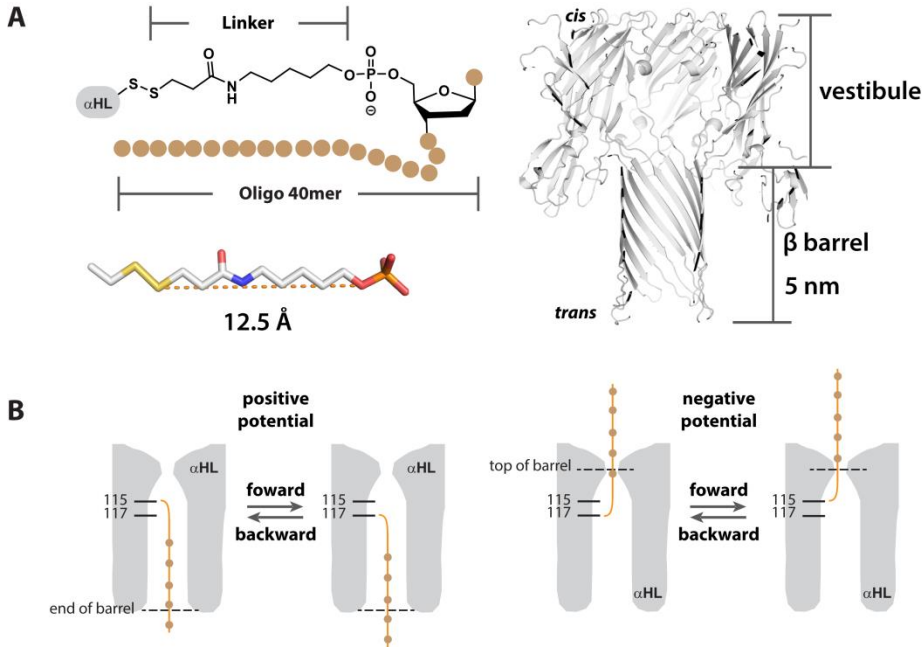


Fig. 3.10. Estimating the number of nucleotides inside the αHL barrel during forward and backward stepping between footholds 115 and 117. (**A**) Length of the fully stretched linker (left) and the αHL β barrel (from residue 111 to 129) (right). Given that the average vertical distance between footholds is 5.6 Å, the linker spans a distance of about two steps. (**B**) Under a positive potential, hopper 1 moved forward from foothold 115 to 117 and backward from 117 to 115 (left). Four negative charges are estimated to be moved within the β barrel during the forward step 115 to 117 (see section 9). Under a negative potential, the hopper moved forward from foothold 117 to 115 and backward from 115 to 117 (right). In this case, there is little charge movement within the β barrel, because most of the DNA has exited into the vestibule.

Table 3.5. Hopping rates (k) and mean dwell times ($\langle\tau\rangle$) derived by QuB[a] for hopper 1 on the two-cysteine track (Footholds: Cys-115, 117) at ±100 mV, ±150 mV, and ±180 mV[b].

	forward stepping			**backward stepping**			
V/mV	$k_{115\text{-}117}$ / s^{-1}	$\langle\tau\rangle_{115\text{-}117}$ / s	**n**	$k_{117\text{-}115}$ / s^{-1}	$\langle\tau\rangle_{117\text{-}115}$ / s	n	$K_{115\text{-}117}$[c]
+100	0.075 ± 0.014	13 ± 3	28	0.055 ± 0.010	18 ± 3	28	1.4 ± 0.4
+150	0.28 ± 0.05	3.5 ± 0.6	40	0.067 ± 0.011	15 ± 2	40	4.2 ± 1.0
+180	0.17 ± 0.03	6.0 ± 1.2	34	0.034 ± 0.006	30 ± 5	34	4.9 ± 1.3
	$k_{117\text{-}115}$ / s^{-1}	$\langle\tau\rangle_{117\text{-}115}$ / s	**n**	$k_{115\text{-}117}$ / s^{-1}	$\langle\tau\rangle_{115\text{-}117}$ / s	n	$K_{117\text{-}115}$[c]
-100	0.49 ± 0.10	2.0 ± 0.4	25	0.018 ± 0.004	55 ± 11	25	27 ± 7
-150	0.66 ± 0.14	1.5 ± 0.3	24	0.025 ± 0.005	39 ± 8	24	26 ± 8
-180	0.89 ± 0.23	1.1 ± 0.3	20	0.022 ± 0.005	45 ± 10	20	40 ± 14

[a] Dwell time analysis and rate constant estimations were performed by using the maximum interval likelihood algorithm of QuB, according to the two-state kinetic models (Fig. 3.11).

[b] Data were collected from four separate αHL pores.

[c] The equilibrium constant $K = k_{forward}/k_{backward}$.

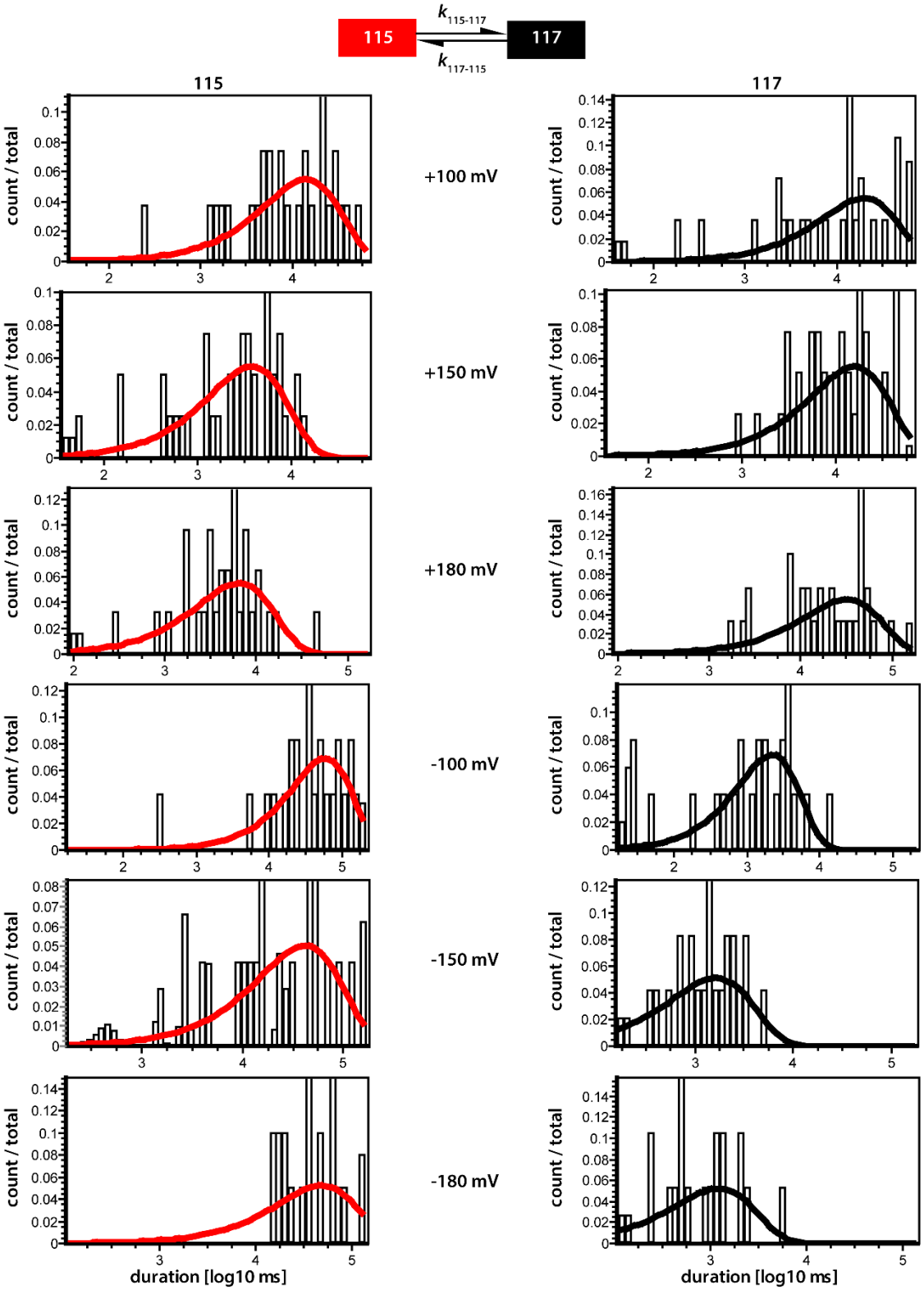


Fig. 3.11. Kinetic model and duration histograms for the forward and backward steps on the two-cysteine track (Footholds: Cys-115, 117) at ±100 mV, ±150 mV, ±180 mV. The forward and backward rate constants (**Table 3.5**) were derived by using the maximum interval likelihood algorithm of QuB, according to the two-state kinetic model (top). Data were collected from four αHL pores.

3.3.3 Hopping rates at ±100 mV, ±150 mV, ±180 mV for the two steps of the three-cysteine track

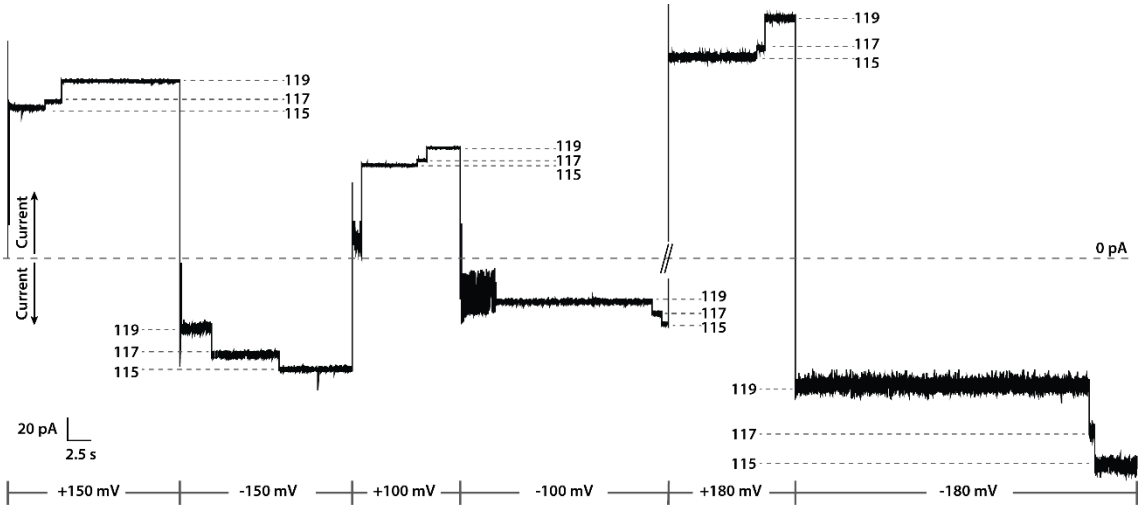


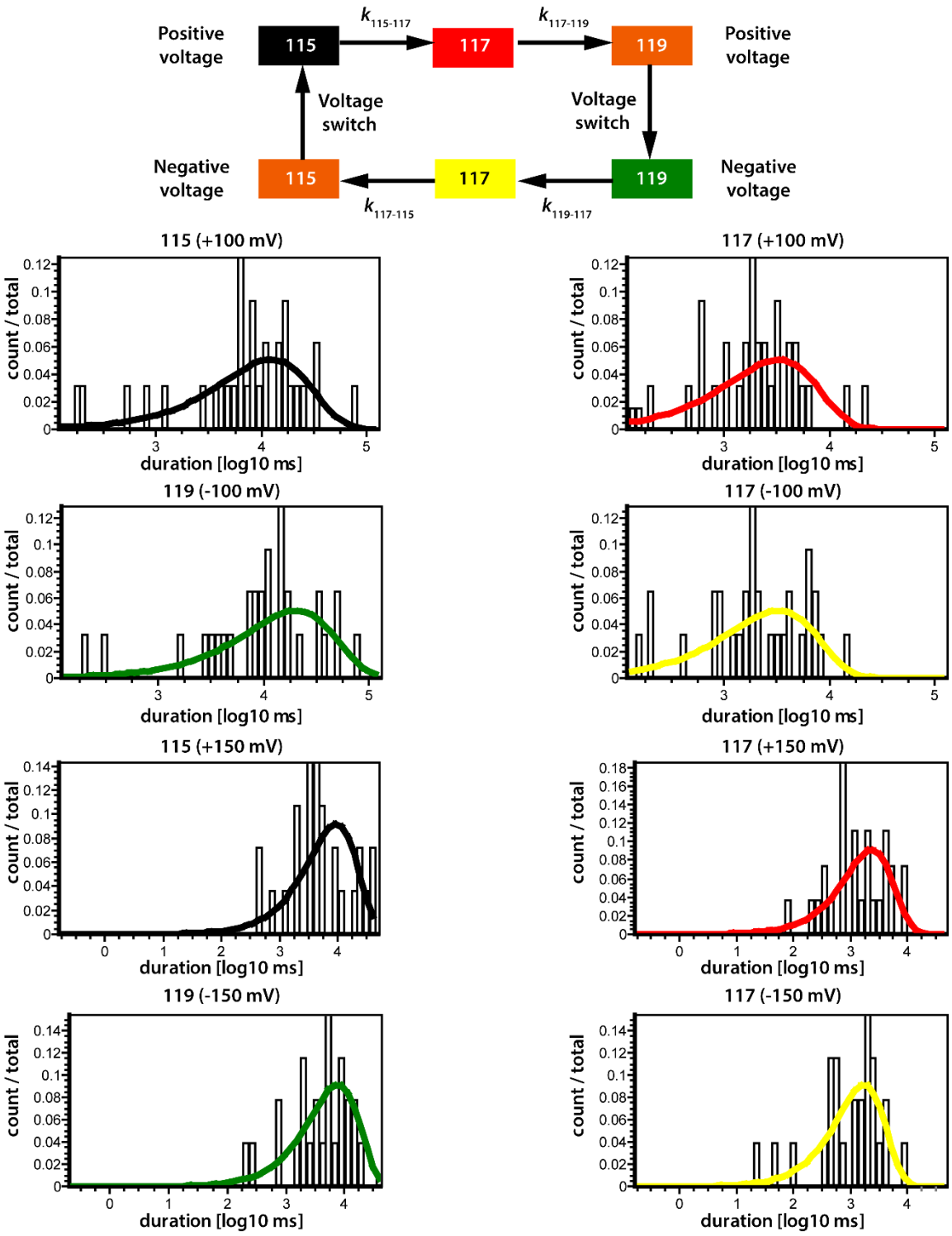
Fig. 3.12. Hopping on a three-cysteine track. With a three-cysteine track (Footholds: Cys-115, 117, 119), two hopping steps were observed at ±100 mV, ±150 mV, ±180 mV with hopper 1. The hopping cycle was repeated >20 times for each voltage and QuB was used to derive the rate constants given in **Table 3.6.** Hopping rates (*k*) and mean dwell times (<τ>) derived by QuB[a] for hopper 1 on the three-cysteine track (Footholds: Cys-115, 117, 119) at ±100 mV, ±150 mV, and ±180 mV[b].. The application of ±100 mV was not sufficient to quickly flip the oligo to the opposite direction and resulted in a very noisy level before the oligo was fully stretched and hopping could re-start. The noisy level was especially common at negative voltages, when it was once observed to last more than 20 min. The trace was filtered at 200 Hz. Conditions: 2 M KCl, 20 mM HEPBS, 20 μM EDTA, pH 8.5, 20 ± 1 °C.

Table 3.6. Hopping rates (k) and mean dwell times ($<\tau>$) derived by QuB[a] for hopper 1 on the three-cysteine track (Footholds: Cys-115, 117, 119) at ±100 mV, ±150 mV, and ±180 mV[b].

	$k_{115\text{-}117}$ / s^{-1}	$<\tau>_{115\text{-}117}$ / s	n	$k_{117\text{-}119}$ / s^{-1}	$<\tau>_{117\text{-}119}$ / s	n
+100 mV	0.086 ± 0.014	12 ± 2	31	0.33 ± 0.05	3.0 ± 0.5	31
+150 mV	0.12 ± 0.02	8.7 ± 1.7	27	0.46 ± 0.09	2.2 ± 0.4	27
+180 mV	0.18 ± 0.03	5.6 ± 1.1	32	0.63 ± 0.11	1.6 ± 0.3	32
	$k_{117\text{-}115}$ / s^{-1}	$<\tau>_{117\text{-}115}$ / s	n	$k_{119\text{-}117}$ / s^{-1}	$<\tau>_{119\text{-}117}$ / s	n
-100 mV	0.33 ± 0.06	3.1 ± 0.6	31	0.053 ± 0.009	19 ± 3	31
-150 mV	0.65 ± 0.13	1.5 ± 0.3	27	0.15 ± 0.03	6.7 ± 1.3	27
-180 mV	0.44 ± 0.09	2.3 ± 0.5	31	0.12 ± 0.02	8.5 ± 1.3	31

[a] Dwell time analysis and rate constant estimations were performed by using the maximum interval likelihood algorithm of QuB. An arbitrary duration of 1 s was assigned to the final foothold in either direction (Cys-119 for cis-to-trans hops at positive voltages or Cys-115 for trans-to-cis hops at negative voltages) to complete the kinetic model (Fig. 3.13).

[b] Data were collected from two αHL pores.


Positive voltage
115
$k_{115\text{-}117}$
117
$k_{117\text{-}119}$
119
Positive voltage
Voltage switch
Voltage switch
Negative voltage
115
117
119
Negative voltage
$k_{117\text{-}115}$
$k_{119\text{-}117}$
115 (+100 mV)
117 (+100 mV)
119 (-100 mV)
117 (-100 mV)
115 (+150 mV)
117 (+150 mV)
119 (-150 mV)
117 (-150 mV)
count / total
duration [log10 ms]

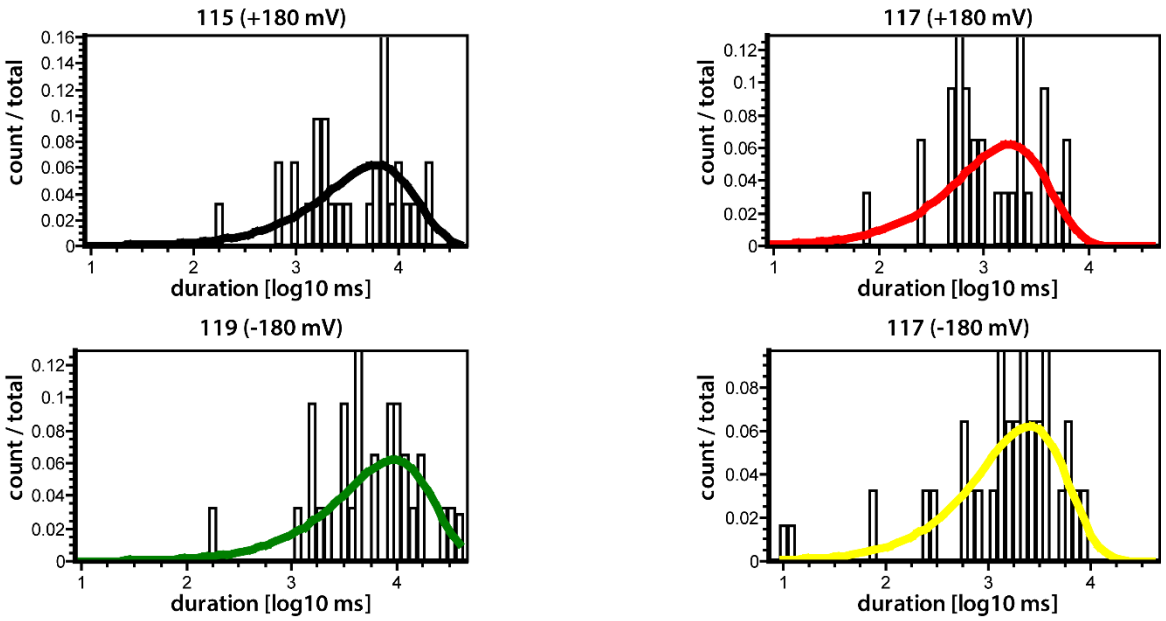


Fig. 3.13. Kinetic model and duration histograms for the two steps on the three-cysteine track (Footholds: Cys-115, 117, 119) at ±100 mV, ±150 mV, ±180 mV. The kinetic model (top) links the two steps at positive voltages with the two at negative voltages to form a closed cycle. Two voltage-switch transitions were built into the model by assigning an arbitrary duration of 1 s to the final foothold (Cys-119 at positive voltages; Cys-115 at negative voltages). Given the absence of back-stepping on the middle foothold during all hopping cycles, QuB was set to treat each forward step as irreversible. QuB derived kinetic rate constants are given in **Table 3.6**. Data were collected from two separate αHL pores.

3.3.4 Effective concentrations of thiol footholds on the cysteine tracks

For a thiol-disulfide interchange:

$$RSSR + R'SH \rightarrow RSSR' + RSH$$

The apparent rate constant k is pH dependent (reaction occurs with the thiolate anion) and is estimated to be ~10 $M^{-1}s^{-1}$ at pH 8.5. The estimation was based on reported rate constants for thiol-disulfide interchanges between L-cysteine, DL-dithiothreitol, L-glutathione and their disulfide forms, for example (*18, 23, 123*).

In the case of an αHL pore, assuming that the disulfide (RSSR) is attached to the protein, and R'SH is in solution, we know:

$$\langle\tau\rangle = 1 / (k\,[R'SH])$$

when R'SH is in solution.

If R'SH is attached to the protein, [R'SH] is an effective concentration in an intramolecular reaction.

For various values of $\langle\tau\rangle$ (see earlier sections), assuming k = 10 $M^{-1}s^{-1}$, we find:

$\langle\tau\rangle$ / s	$[RSH]_{effective}$ / mM
1	100
5	20
20	5

The effective concentration of a thiol on the track with respect to an upstream disulfide is then in the mM range. Effective concentrations in this range were found for the second arm of a chelator in a nanoreactor (*65*) and for the intramolecular binding of an enzyme inhibitor (*124*).

3.4 Energetics of the directional movement

In the case of forward stepping under +150 mV from foothold 115 to 117 on a two-cysteine track (Footholds: Cys-115, 117) (as discussed in section 3.3.2), the force imparted on the hopper-oligo is

$$F = n(qe)E$$

n = number of phosphodiesters in the field

q = fractional charge on a phosphodiester

e = elementary charge (1.60×10^{-19} C)

E = electric field

Previously, it was suggested that the applied transmembrane potential drops mostly across the transmembrane β barrel (5 nm) of the αHL pore (*56*). At ±150 mV, E is therefore 0.03 V nm^{-1}.

Let n = 4 (Fig. 3.10B) and q = 1

$F = n(qe)E = 4 \times 1.60\times10^{-19}$ C × 0.03 V $nm^{-1} = 1.92\times10^{-11}$ N = 19.2 pN

Let n = 4 and q = 0.4 (due to shielding effect of high salt (*125, 126*))

$F = n(qe)E = 4 \times 0.4 \times 1.60\times10^{-19}$ C × 0.03 V $nm^{-1} = 7.68\times10^{-12}$ N = 7.68 pN

The work done by the electric field W = Fd per molecule = FdN_A per mole (d = 0.56 nm)

q = 1, $W = 1.08\times10^{-20}$ J = 6.47 kJ mol^{-1}

q = 0.4, $W = 4.32\times10^{-21}$ J = 2.59 kJ mol^{-1}

Similarly, the force applied and the work done to move the hopper-oligo complex from 117 to 115 under -150 mV can be calculated with n = 1 (Fig. 3.10B).

q = 1, F = 4.80 pN, W = 1.62 kJ mol^{-1}

q = 0.4, F = 1.92 pN, W = 0.648 kJ mol^{-1}

From the experimentally determined equilibrium constants ($K = k_{forward}/k_{backward}$) (Table 3.5), the corresponding ΔG can be calculated following $\Delta G = -RT\ln K$ (R = 8.314 $JK^{-1}mol^{-1}$, T = 293.15 K).

	115-117 (+150 mV)	**117-115 (-150 mV)**
K	4.2 ± 1.0	26 ± 8
ΔG (kJ mol^{-1})	-3.5	-7.9
-W (kJ mol^{-1}) (q = 0.4)	-2.59	-0.648

The difference between the derived ΔG and the calculated W can be accounted for by other variable thermodynamic terms (either positive or negative), which include the aforementioned DNA coiling in the pore's vestibule. After all, the system is highly complex with 1) non-uniform side-chains; 2) non-uniform cross section of the β barrel; 3) non-uniform threaded polymers.

The non-electrostatic terms will be more important to the equilibrium for each step at low applied voltages. However, processive multistep hopping only requires that the overall equilibrium constant ($K_{overall}$) is larger than 1, which can be calculated as the product of the equilibrium constants for each step (e.g. $K_{overall} = K_1K_2K_3K_4$). For example, using the experimentally derived $K_{115\text{-}117}$ at +150 mV for all four steps on a five-cysteine track,

$$K_{overall} = K^4 = (4.2)^4 = 311$$

This large $K_{overall}$ explains the observed processive and directional hopping.

It is worth pointing out that the calculations shown here only apply to stepping between the footholds 115 and 117. All thermodynamic contributors, including the electrostatic work, are likely to vary between footholds. For instance, the charged oligo moves stepwise out of the electric field as the hopper steps forward along the track. Consequently, the electrostatic contribution to ΔG becomes smaller as the hopper approaches the terminal foothold.

3.5 Releasing the hopper from a predefined foothold

Each subunit of the αHL pore offers two antiparallel β strands to the transmembrane β barrel with an inter-strand distance of ~5 Å (Cα-Cα). Given that the formation of cross-strand disulfides has been reported (*127*), we reasoned that the addition of a cysteine on an adjacent strand would compel hopper release from the track at a designated foothold. Indeed, with an L-shaped track consisting of cysteines at positions 115, 117, 119, and 139, the hopper attached to the track at foothold 115 by regioselective disulfide formation and dissociated from the track when it reached foothold 119. The release was initiated by Cys-139, through thiol-disulfide interchange to form a cross-strand disulfide bridge, which blocked the access of subsequent hoppers to foothold 119 (Fig. 3.14). The preference for hopper release versus hopper transfer to the adjacent strand is attributed to an unfavourable collinear alignment of the three participating sulfur atoms necessary for transfer. In the future, the engineering of footholds on a surface will allow the construction of more complex hopping pathways where hoppers are transferred to new tracks at designed junctions and cargos are released at predesignated depots.

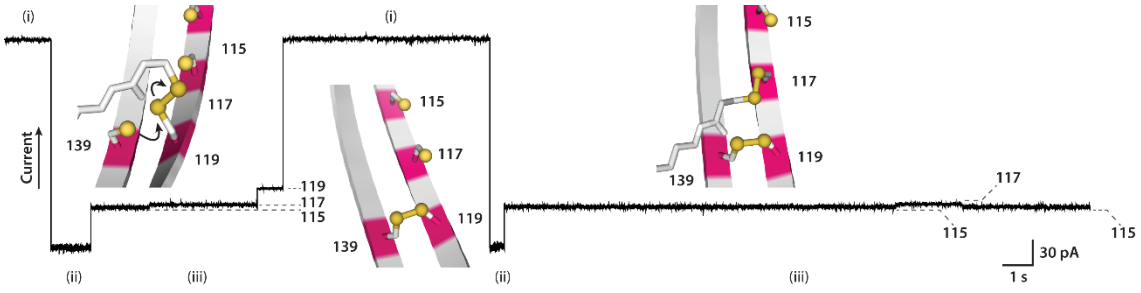


Fig. 3.14. Releasing the hopper from a predefined foothold. On an L-shaped track consisting of cysteines at positions 115, 117, 119, and 139, the hopper moved along the track from Cys-115 to Cys-119, where it was released by the side chain of Cys-139. Subsequently, a second hopper became loaded at Cys-115, but its motion was arrested at Cys-117 because Cys-119 was now engaged in an interstrand disulfide bond [(i), (ii), (iii) as in Fig. 3.3B]. Backstepping to Cys-115 was also seen. Conditions: 2 M KCl, 20 mM HEPBS, 20 μM EDTA, pH 8.5, 20° ± 1°C.

3.6 Oligonucleotide sequencing

Current nanopore sequencing technology uses a biological machine—an unwinding enzyme—to ratchet a single-stranded DNA base-by-base through a protein nanopore (Fig. 3.15). Here, the ability to translocate a stretched DNA cargo while maintaining a covalent bond with the nanopore suggests a method for the chemical ratcheting of a nucleic acid during nanopore sequencing (*128*), which was explored in a proof-of-concept experiment. To provide a marker, two adjacent abasic residues (1′,2′-dideoxyribose, dS) (*129*) were incorporated into the cargo oligos carried by hoppers 2 and 3, and current patterns were recorded during four-step hopping between Cys-113 and Cys-121 (Fig. 3.16).

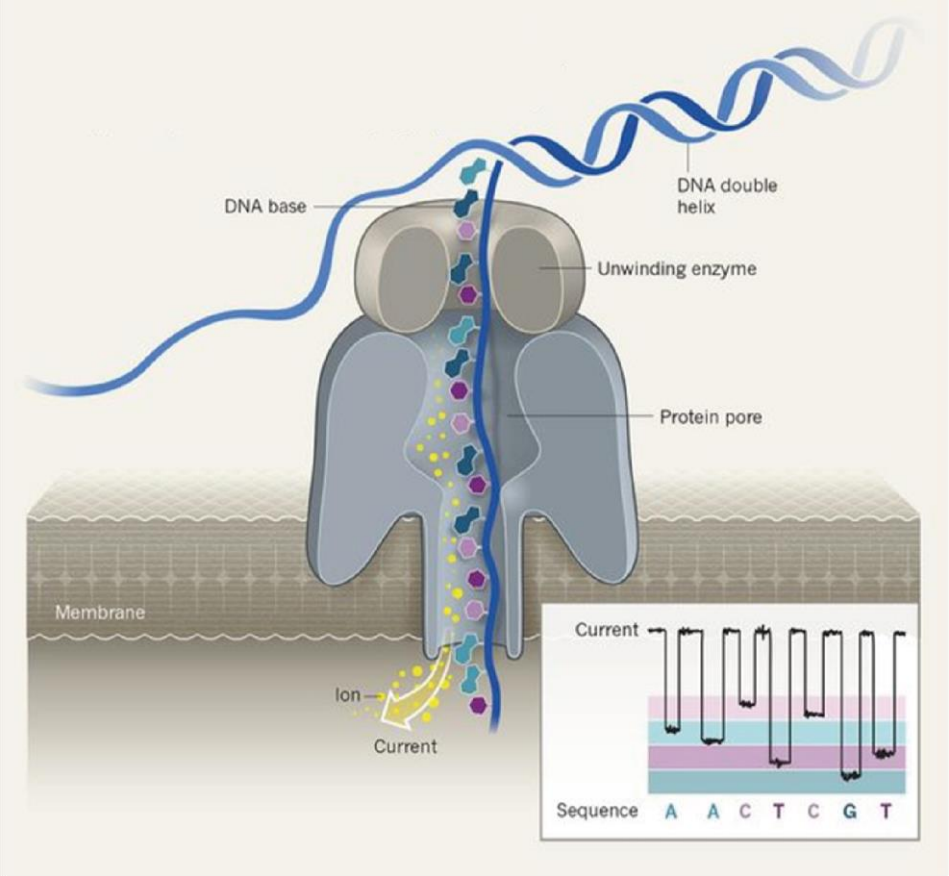


Fig. 3.15. Basis of nanopore sequencing. An enzyme unwinds DNA, ratcheting one strand base-by-base through a protein pore. The sequence is determined by analysis of the fluctuations of the ionic current flowing through the nanopore. (Image produced by Nik Spencer/Nature)

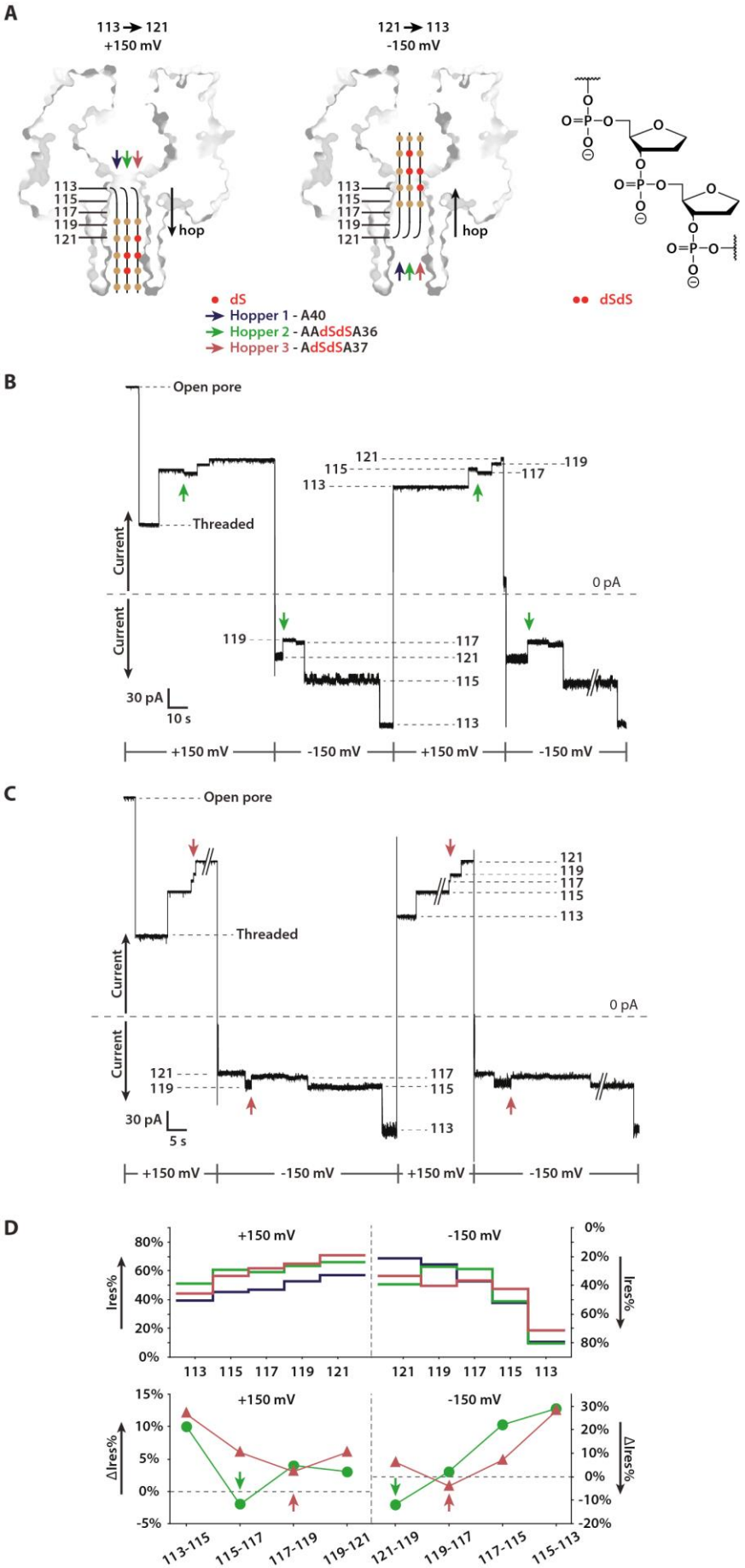

A
113 → 121
+150 mV
121 → 113
-150 mV
113
115
117
119
121
hop
dS
dSdS
Hopper 1 - A40
Hopper 2 - AAdSdSA36
Hopper 3 - AdSdSA37
B
Open pore
Threaded
Current
0 pA
30 pA
10 s
+150 mV
-150 mV
C
5 s
D
+150 mV
-150 mV
Ires%
ΔIres%
113-115
115-117
117-119
119-121
121-119
119-117
117-115
115-113

Fig. 3.16. **Discrimination of different DNA cargos.** (**A**) Sequence design of DNA cargos. Two sequential abasic nucleotides (dSdS) were substituted at positions 3 and (hopper 2) or at positions 2 and 3 (hopper 3). The numbers of nucleotides (brown circles, dA; red circles, dS) placed inside the β barrel are based on PyMOL modelling. (**B**) With a five-cysteine track, four-step hopping was observed with hopper 2 at ±150 mV. The current decreases for hops from 115 to 117 and from 121 to 119 are marked (green arrows). (**C**) Four-step hopping with hopper 3 at ±150 mV. The current transitions for hops from 117 to 119 and from 119 to 117 are marked (red arrows). (**D**) Top: Overlaid current traces of hoppers 1, 2, and 3 with step durations normalized. The current levels are given as the residual current with respect to the open pore level (Ires%). Bottom: Step sizes of hoppers 2 and 3 plotted as ΔIres%. Minima in the plots showing the single nucleotide offset are marked. Conditions: 2 M KCl, 20 mM HEPBS, 20 μM EDTA, pH 8.5, 20° ± 1°C.

By comparison with hopper 1, hoppers 2 and 3 showed different patterns of current modulation (Table 3.7, Fig. 3.5A, Fig. 3.16, A and C, Fig. 3.17). The conductance patterns generated by the four-step hopping motion could be repeated with different molecules of hopper 2 and hopper 3 (n = 3 for each hopper), establishing the patterns as clear identifiers of each cargo sequence. The residual currents (Ires%, the remaining current as a percentage of the open pore current) for the three hoppers residing at each foothold were plotted for comparison (Fig. 3.16C). Hopper 1 and hopper 2 gave almost identical current blockades at each of footholds 115 and 113 under -150 mV (Fig. 3.16C), implying that the dSdS sequence had been transported well out of the sensing region by hopper 2. Moreover, hopper 2 and hopper 3 have a single nucleotide offset in the dSdS positions and we observed a one-step offset between hopper 2 and hopper 3 in ΔIres%, the difference in Ires% between two successive steps (Fig. 3.16C; the vertical step size, 5.6 Å, is similar to the inter-nucleotide distance in stretched single-stranded DNA, 6.9 Å (*75*)). These observations demonstrate that the hopper system reported here has the potential to discriminate bases for sequencing purposes (*129*). An advantage of a processive hopper, which might

improve sequencing accuracy, is the ability to reverse the chemical ratcheting process and thereby obtain many-fold coverage of an individual DNA strand. Of course, the present system is limited by its short track. While longer β-barrel pores exist (*130*, *131*), a viable sequencing process will require protracted ratcheting of numerous DNA strands in parallel, perhaps by using footholds on an extended crystalline surface or internal thiophosphate feet to transport long replica strands over relatively short tracks.

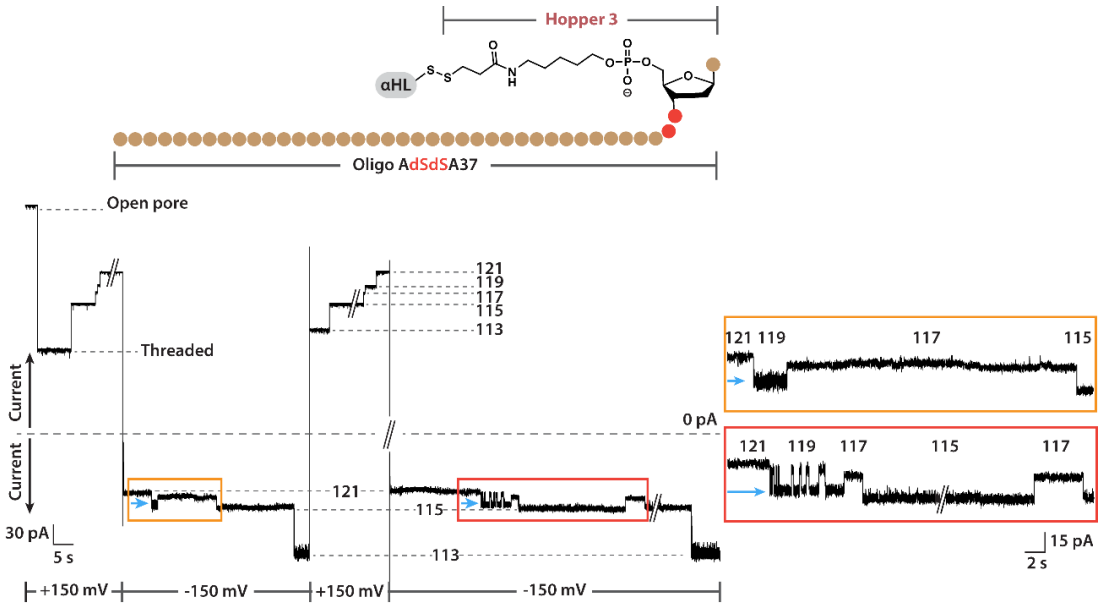


Fig. 3.17. **Two current patterns for hopper 3 on foothold 119 at -150 mV.** Four-step hopping was observed with hopper 3 at ±150 mV (n = 3 separate experiments). While the hopping generated a consistent pattern at +150 mV, hopping at -150 mV produced two ionic current signatures. In one, the current at foothold 119 was significantly noisier than the current with the hopper at neighbouring footholds (top expanded trace): root mean square noise $I_{RMS}(119) = 2.8 \pm 0.1$ pA; $I_{RMS}(121) = 1.2 \pm 0.2$ pA; $I_{RMS}(117) = 1.0 \pm 0.1$ pA. In the second pattern, the current with hopper 3 at foothold 119 displayed two sub-conductance levels (bottom expanded trace). The interconversion between two patterns was not observed. The ratio of dwells at foothold 119 with the noisy appearance to dwells with the two sub-conductance states was 5:2. The current amplitude is the same for the noisy level of the first pattern and the higher conductance state (larger negative current) in the second pattern (indicated with blue arrows). This value was used to plot **Fig. 3.16**C and was reported in **Table 3.7**. The two patterns must correspond to two modes of interaction of hopper 3 with the nanopore, as they were seen not only in separate experiments, but also with the same hopper molecule in the same experiment, which excludes the possibilities of an impurity in the oligo or hopper-to-hopper variation.

Table 3.7. Residual current (Ires%) pattern for each hopper on the five-cysteine track (Footholds: Cys-113, 115, 117, 119, 121).

+150 mV	Ires%[a]			-150 mV	Ires%[a]		
	Hopper 1	**Hopper 2**	**Hopper 3**		**Hopper 1**	**Hopper 2**	**Hopper 3**
Threading	33%	35%	35%	**/**	/	/	/
113	39%	51%	44%	**113**	80%	80%	71%
115	45%	61%	56%	**115**	52%	51%	43%
117	47%	59%	62%	**117**	37%	29%	36%
119	53%	63%	65%	**119**	25%	27%	40%
121	57%	66%	71%	**121**	21%	39%	34%

[a] The standard deviations of Ires% are less than 0.5% in all cases (derived from n = 3 separate experiments for each hopper).

3.7 Supplementary materials and methods

3.7.1 Synthesis and characterization of hopper-carrier conjugates

3.7.1.1 Chemicals

Acetic anhydride (Ac_2O), acetonitrile (HPLC grade), 5-azidopentanoic acid, N,N-diisopropylethylamine (DIPEA), diethyl ether, dimethylformamide (DMF), hexafluoro-2-propanol (HFIP), N,N-hydroxybenzotriazole (HOBt), 1-methyl-2-pyrrolidinone (NMP), Nα-Fmoc-Nε-biotinyl-L-lysine, piperidine, triethylamine (TEA), and trifluoroacetic acid (TFA) were purchased from Sigma-Aldrich. 2-(1H-Benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU), Fmoc-Phe-OH, and Rink Amide MBHA resin LL (100-200 mesh) were from Novabiochem. Fmoc-NH-PEG(12)-COOH was purchased from Iris Biotech GmbH.

3.7.1.2 Carrier preparation using solid-phase peptide synthesis

The carrier (Fig. 3.18A) was made by using manual solid-phase peptide synthesis, adapted from previously reported protocols (*67*).

Peptide elongation was carried out on 0.25 mmol Rink Amide MBHA resin LL (0.37 mmol g^{-1}) by the Fmoc method with HBTU and DIPEA as coupling reagents. The resin was washed and swollen overnight in DMF and Fmoc was removed by treatment with 20% (v/v) piperidine in NMP at room temperature. The carrier was assembled through standard chain elongation by using the reactants Nα-Fmoc-Nε-biotinyl-L-lysine, Fmoc-Phe-OH, Fmoc-NH-PEG(12)-COOH, and 5-azidopentanoic acid. For each of the four coupling reactions, a solution of 1.0 mmol reactant, 0.95 mmol HBTU, and 2.0 mmol DIPEA in 1.9 mL DMF was added to the resin, which was then shaken for 12 min. After each coupling, the resin was washed with NMP and capped with 1.5 mmol Ac_2O, 0.044 mmol HOBt, and 0.39 mmol DIPEA in 3 mL DMF for 2 min. The capped

resin was then washed with NMP again before the next coupling. After the completion of chain assembly, the resin was washed sequentially with DMF and ethanol and dried *in vacuo*. Peptides were cleaved from the resin by TFA treatment for 2 h. The TFA was evaporated in a stream of nitrogen and the peptide was precipitated with cold diethyl ether, followed by centrifugation, and several trituration treatments with diethyl ether.

3.7.1.3 Carrier purification and characterization

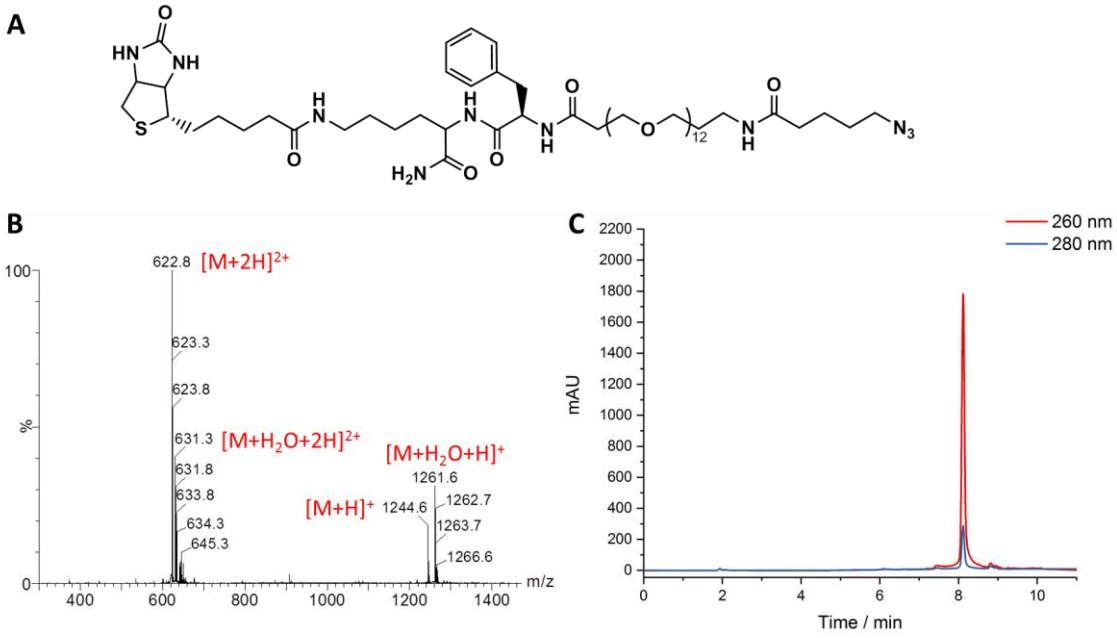


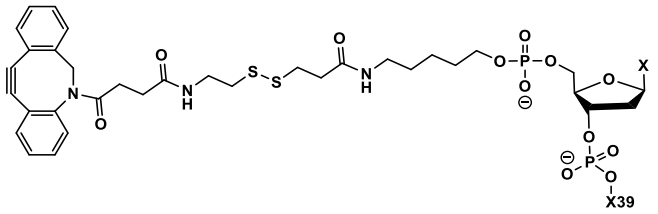
Fig. 3.18. LC-MS and HPLC characterization of the carrier. (**A**) Structure of the carrier. (**B**) The carrier mass ($[M+H]^+$) was calculated to be 1243.7 g mol^{-1} and found to be 1244.6 g mol^{-1}. (**C**) The purity of the carrier was confirmed by analytical HPLC.

The crude peptide was dissolved in DMF and purified by preparative RP-HPLC (Dionex UltiMate 3000; Vydac C18 column: 250 × 22 mm, 10-15 µm (particle size); linear gradient: 5-95% eluant B in eluant A over 45 min, flow rate: 15 mL min^{-1}; eluant A: 0.1% TFA in water; eluant B: 0.1% TFA in acetonitrile). The identity and purity of the fractions were determined by analytical LC-MS (Waters LCT accurate-mass time-of-flight instrument (ESI-positive); Chromolith RP-18e column: 50 mm × 2 mm; linear gradient: 5-100% eluant B in eluant A over 8 min, flow rate: 1 mL min^{-1}; eluant A: 0.1%

formic acid in water; eluant B: 0.1% formic acid in acetonitrile). The fractions showing the correct mass were combined, lyophilized, and characterized by LC-MS and analytical RP-HPLC (Agilent 1260 Infinity HPLC; Polaris C18 column: 150 × 4.6 mm, 5 μm (particle size); linear gradient: 5-95% eluant B in eluant A over 10 min, flow rate: 1 mL min^{-1}; eluant A: 0.1% TFA in water; eluant B: 0.1% TFA in acetonitrile).

3.7.1.4 Generation and characterization of hopper-carrier conjugates

Oligonucleotides modified with dibenzocyclooctyne-disulfide (DBCO-SS) at the 5' terminus were purchased from biomers.net.



5’ DBCO-SS modification (x = A or dS)

5' DBCO-SS-A40 oligo (10 μL, 1 mM in MilliQ water) was added to the carrier (1 μL, 10 mM in MilliQ water). The mixture was left at room temperature for 30 min to form the hopper-1-carrier conjugate by a copper-free click reaction. Similarly, the hopper-2-carrier and hopper-3-carrier conjugates were assembled with 5' DBCO-SS-AAdSdSA36 and 5' DBCO-SS-AdSdSA37, respectively.

The conjugates were purified by semi-preparative HPLC (Agilent 1260 Infinity; Supelco Discovery BIO Wide Pore C18 column: 250 × 10mm, 10 μm (particle size); linear gradient: 10-90% eluant B in eluant A over 30 min, flow rate: 4.5 mL min^{-1}; eluant A: 0.1% TFA in water; eluant B: 0.1% TFA in acetonitrile) and identified by LC-MS (UPLC-MS Waters XEVO G2-QTOF (ESI-negative); ACQUITY UPLC Oligonucleotide BEH C18 Column: 2.1 mm X 50 mm, 1.7 μm (particle size); linear gradient: 0-70% eluant B in eluant A over 8 min, flow rate: 0.2 mL min^{-1}; eluant A: 8.6

mM TEA, 200 mM HFIP in 5% methanol/water (v/v); eluant B: 20% eluant A in methanol).

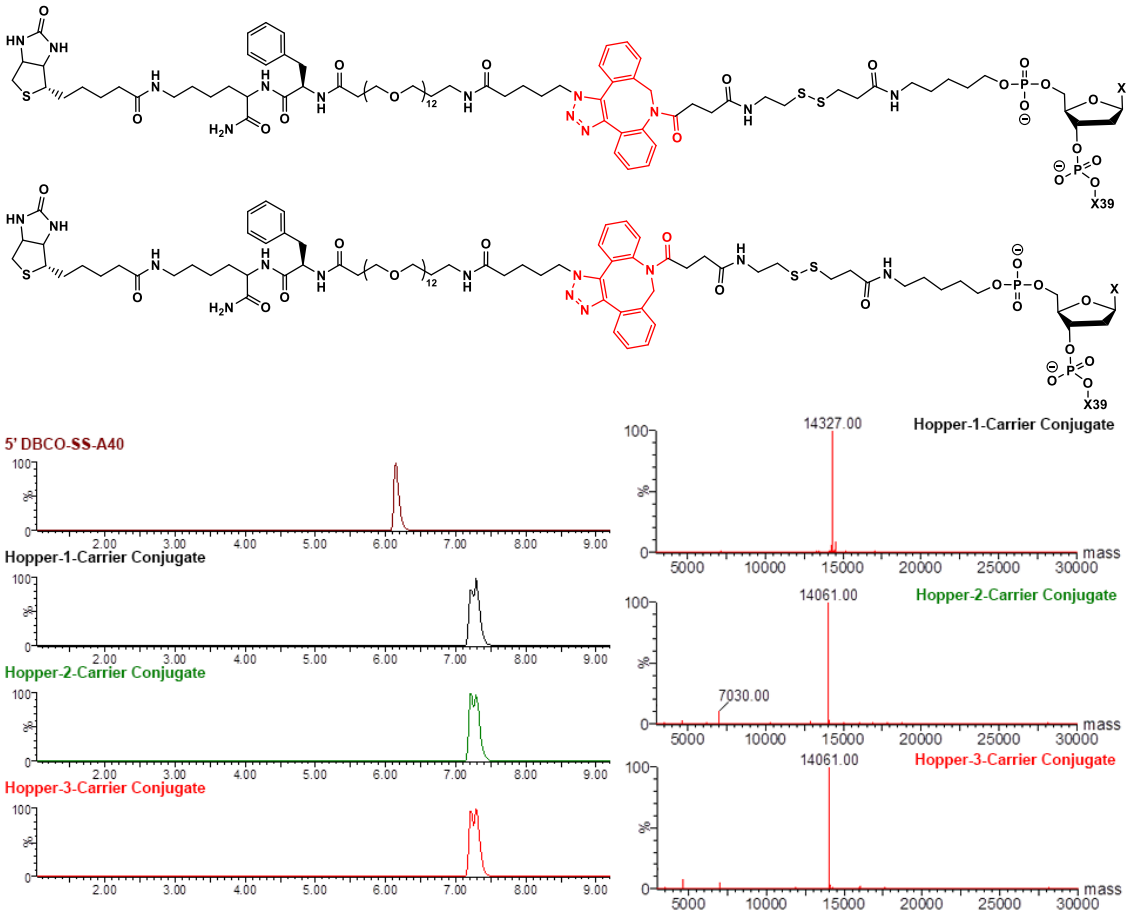


Fig. 3.19. LC-MS characterization of hopper-carrier conjugates. The two peaks in each chromatogram have the same mass, corresponding to the two regioisomeric conjugates (top) formed by the copper-free click reaction. The mass of hopper-1-carrier conjugate was calculated to be 14327 g mol^{-1} and found to be 14327 g mol^{-1}. The masses of hopper-2-carrier conjugate and hopper-3-carrier conjugate were both calculated to be 14059 g mol^{-1} and both found to be 14061 g mol^{-1}.

3.7.2 Preparation of αHL monomers with multi-cysteine tracks

The construction of pT7–αHL–115C117C-D8H6 and pT7–αHL–113C115C117C119C121C-D8H6 (pT7–αHL-5C-D8H6), which encode αHL mutants containing cysteines either at positions 115 and 117, or at positions 113, 115, 117, 119,

and 121, as well as an octa-aspartate tail and a His-tag at the C terminus, have been previously reported by our group (*91*).

pT7–αHL–6C-D8H6 encodes a αHL mutant containing cysteines at positions 113, 115, 117, 119, 121, and 123 as well as an octa-aspartate tail and a His-tag at the C terminus. It was generated from pT7–αHL-5C-D8H6 by *in vivo* homologous recombination. Two sets of PCR reactions were carried out. For the first reaction, the template was first linearized by NdeI (New England Biolabs) before PCR with the forward mutagenic primer: 5'-CTGCTTCTGCGGTTGTGTTACTGGTGATGATACAGG-3' and the reverse non-mutagenic primer (SC47): 5′-CAGAAGTGGTCCTGCAACTTTAT-3′. For the second reaction, the template was linearized by HindIII (New England Biolabs) before PCR with the reverse mutagenic primer: 5'-CCTGTATCATCACCAGTAACACAACCGCAGAAGCAG-3' and the forward non-mutagenic primer (SC46): 5′-ATAAAGTTGCAGGACCACTTCTG-3′. The linearized templates and primers were mixed with 1XPhusion Flash HF Mastermix (New England Biolabs) and put through the following cycling program: 94°C for 5 min, then 30 cycles of 94°C (30 s), 50°C (30 s), 72°C (30 s), and finally 50°C for 5 min. After PCR, 5 μL of each reaction were mixed and used to transform XL10-Gold cells (Agilent). The transformed cells were grown at 37°C overnight on LB (Luria Broth)-carbenicillin plates. Plasmid DNA was isolated from colonies by using the QIAprep Spin Miniprep Kit (QIAGEN). Successful mutagenesis was confirmed by DNA sequencing.

Similarly, pT7–αHL–115C117C119C-D8H6 was made from pT7–αHL-WT-D8H6 with the mutagenic primer pair: 5'-CAAAAGAGTATATGAGTTGCTTATGCTATTGCTTCAACG-3' (Forward); 5'-CGTTGAAGCAATAGCATAAGCAACTCATATACTCTTTTG-3' (Reverse). pT7–αHL–115C117C119C139C-D8H6 was made from pT7–αHL–115C117C119C-D8H6 with the mutagenic primer pair: 5'-

CCTTATTGGTGCATGTGTTTCGATTGGTCATACACTG-3' (Forward); 5'-CAGTGTATGACCAATCGAAACACATGCACCAATAAGG-3' (Reverse).

3.7.3 Preparation of an αHL heptamer containing a multi-cysteine tracks on one of the seven subunits

Engineered αHL polypeptides were expressed by using a commercial *in vitro* transcription-translation (IVTT) kit: *E. coli* T7 S30 Extract System for Circular DNA (Promega). To suppress transcription by *E. coli* RNA polymerase, the T7 S30 extract provided in the kit was treated with rifampicin prior to use (1 μg mL^{-1}, final concentration). A standard reaction comprised: DNA template (3.2 μg), amino acid mix minus methionine (supplied with the kit, 5 μL), S30 premix without amino acids (supplied with the kit, 20 μL), [^{35}S]methionine (2 μL, 1,200 Ci $mmol^{-1}$, 15 mCi mL^{-1}, MP Biomedicals), 15 μL T7 S30 extract (supplied with the kit. 15 μL), and nuclease-free water to a final volume of 50 μL. To make heteroheptamers, plasmids encoding the WT αHL and the mutant αHL were mixed in a ratio 6:1 (WT:mutant). The IVTT mixture was incubated at 37°C for 1 h.

Heptamerization was carried out by the addition of rabbit erythrocyte membranes (3 μL, ~1 mg protein mL^{-1}) to the IVTT reaction mixture (50 μL), followed by incubation at 37°C for another 1 h. The mixture was then centrifuged for 10 min at 25,000 × g. The supernatant was removed and the pellet resuspended in MBSA buffer (200 μL, 10 mM 3-morpholinopropane-1-sulfonic acid (MOPS), 150 mM NaCl, 1 mg mL^{-1} bovine serum albumin, pH 7.4). The wash with MBSA was repeated before the pellet was resuspended in 2X Laemmli sample buffer (50 μL) without heating and electrophoresed in a 5% SDS polyacrylamide gel at 70 V for 15 h.

The αHL heteroheptamers containing different numbers of mutant subunits were separated in the gel based on their different electrophoretic mobilities, which were

determined by the number of octa-aspartate (D8) tails (*92*). The top and bottom bands corresponded to WT_7 and $(mutant\text{-}D8H6)_7$, respectively. The second band from the top was the desired heteroheptamer containing a single mutant subunit.

To extract heptameric pores, the gel was first dried without fixation on Whatman 3M filter paper under vacuum for 5 h at room temperature. After visualization by autoradiography with Kodak BioMax MR film, the desired bands were cut from the gel with a scalpel. Each excised band was rehydrated in TE buffer (300 µL, 10 mM Tris·HCl, 1 mM ethylenediaminetetraacetic acid (EDTA), pH 8.0) for 1 h at room temperature. The filter paper was then removed, and the rehydrated gel was macerated with a pestle. The resulting slurry was filtered through a 0.2 µm hydrophilic membrane filter (Proteus Mini Clarification Spin Column, Generon). The filtrate was stored in 10 µL aliquots at -80°C.

3.7.4 Single channel recordings

3.7.4.1 General

1,2-Diphytanoyl-sn-glycerol-3-phosphocholine (DPhPC) was purchased from Avanti Polar Lipids. Unless otherwise stated, the other chemicals were purchased from Sigma-Aldrich.

Planar bilayer recordings were performed following the method established by Montal and Mueller (*93*). Two Delrin compartments were separated by a 25 µm-thick Teflon film containing an aperture (60 µm in diameter). The aperture was pre-treated with 1% (v/v) hexadecane in pentane. Each compartment was then filled with buffer (500 µL, 2 M KCl, 20 mM N-(2-hydroxyethyl)piperazine-N'-(4-butanesulfonic acid) (HEPBS), 20 µM EDTA, pH 8.5). A drop of DPhPC in pentane (5 mg mL^{-1}) was added to each compartment. Repetitive up-and-down pipetting of the buffer solution resulted in the formation of a lipid bilayer across the aperture. A transbilayer potential was

applied with two Ag/AgCl electrodes, each contained within a salt bridge formed from 3 M KCl in 3% (w/v) low-melt agarose.

Ionic currents were recorded by using a patch clamp amplifier (Axopatch 200B, Axon Instruments) with a 4-pole low-pass Bessel filter (80 dB/decade) at room temperature (20 ± 1 °C). Signals were digitized with a Digidata 1320A digitizer (Molecular Devices), connected to a computer running the pCLAMP 10.3 software suite (Molecular Devices). Unless stated otherwise, the signal was filtered with a corner frequency of 5 kHz and sampled at 25 kHz.

3.7.4.2 Monitoring the hopping processes

The hopper-carrier conjugate (400 μM in 1 μL MilliQ water) was added to a solution of traptavidin (Kerafast) (40 μM in 10 μL phosphate-buffered saline, pH 7.4) and incubated at room temperature (20 ± 1 °C) for 15 min to form the traptavidin-tagged hopper-carrier. The cis compartment of the apparatus, containing an engineered αHL construct in 2 M KCl, 20 mM HEPBS, 20 μM EDTA, pH 8.5 (500 μL), was stirred until a pore inserted into the bilayer. The traptavidin-tagged hopper-carrier conjugate (3 μL of a 36 μM solution) was then added to the cis compartment and driven into the pore with a transmembrane potential of +150 mV (cis at ground). In the case of hopper 1, threading of the conjugate reduced the current to Ires% = 33 ± 1% (n = 20). Subsequent loading of hopper 1 onto Cys-115 by disulfide interchange gave a current increase to Ires% = 45 ± 1% (n = 20). Under a positive potential (e.g. +150 mV), the hopper moved from Cys-115 towards the trans end of the track (downward). When the hopper reached the final foothold, a negative potential (e.g. -150 mV) drove the hopper back towards the cis end of the track (upward). Alternation between positive and negative potentials repeatedly drove the hopper back and forth on the track.

We define cis to trans movement as downward and trans to cis as upward. Under a positive potential (e.g. +150 mV), the forward stepping is downward; the backward stepping is upward. Under a negative potential (e.g. -150 mV), the forward stepping is upward; the backward stepping is downward.

3.7.4.3 Single-channel data analysis

Current traces were idealized by using Clampfit 10.3 (Molecular Devices). The idealized data were analyzed with QuB 2.0 software (www.qub.buffalo.edu) (*94*). Dwell time analysis and rate constant estimations were performed by using the maximum interval likelihood (MIL) algorithm of QuB (*76*). In the cases of hopping on a three-cysteine track and a five-cysteine track, an arbitrary duration of 1 s was assigned to the final foothold in either direction (cis-trans or trans-cis) to complete the kinetic models comprising steps at both positive and negative voltages (Fig. 3.8 and Fig. 3.13). A two-state kinetic model was used for hopping between two footholds (Fig. 3.9 and Fig. 3.11).

Chapter 4 Conclusions and Perspectives

4.1 Conclusions

Herein, I summarise the major findings presented in this thesis.

Chapter 2 describes an artificial system that mediates precise selective control over thiol-disulfide interchange. The influence of spatial alignment was demonstrated on the reaction between a nanoreactor cysteine and a disulfide consisting of chemically equivalent sulfur atoms (S_α and S_β). By manipulating the relative position of the substrate disulfide and the thiolate within single-cysteine nanoreactors, we steered the interchange between complete S_β regioselectivity and >80% S_α regioselectivity. Kinetic analysis revealed that the complete S_β regioselectivity was achieved with a smaller rate of thiol-disulfide interchange, in comparison to the cases where reactions occurred towards both S_α and S_β. Furthermore, we constructed nanoreactors containing two cysteines in which two consecutive thiol-disulfide interchanges occurred. Interestingly, when the two cysteines were adjacent on the same β strand (Cys-115 and Cys-117, Cα-Cα inter-cysteine distance ~7 Å), complete regioselectivity was observed as the substrate thiol was passed from one residue to the other. When the two cysteines were located on antiparallel strands (Cys-115 and Cys-143, Cα-Cα inter-cysteine distance ~5 Å), in addition to the substrate transfer between two cysteines, we observed the formation of a cross-strand disulfide as an alternative regioselective outcome. Complete site-selective thiol-disulfide interchange was achieved with nanoreactor 115/117, in which the substrate selectively reacted with Cys-115 in the presence of Cys-117. Therefore, we were able to manipulate the selectivity of thiol-disulfide interchange without altering the intrinsic chemistry of the reactants. The control that was achieved

exploited mandatory collinear arrangement of the three S atoms in the transition state of the reversible S_N2 reaction.

Chapter 3 describes the further development of the chemistry in **Chapter 2** into a molecular hopper, which carries an oligonucleotide cargo while moving directionally and processively along a cysteine track built within the αHL nanopore. Transmembrane potentials of opposite polarities could be applied to drive the hopping in either cis-to-trans or trans-to-cis direction. Base-by-base stepwise movement of an oligonucleotide inside a protein sensor is the pre-requisite for nanopore sequencing technology. In a proof-of-concept experiment, we identified the abasic residues incorporated at different loci within the cargo oligonucleotide by their fingerprints on the current traces during hopping.

4.2 Perspectives and future directions

In this section, I discuss several points of interest, which have not yet been explored and call for future study.

4.2.1 Chemical properties of the constituent S atoms in the disulfide

As mentioned in **Chapter 2**, a chemical reaction often 'innately' prefers one reactive position over the others within a functional group because of their intrinsic difference in chemical properties. In the case of a small-molecule thiol-disulfide interchange, innate regioselectivity prefers the release of the thiol with lower pK_a because it is a better leaving group (*22*). To dissect the influence of spatial alignment from that of innate chemical reactivity on regioselectivity (**Chapter 2**), we used a disulfide consisting of two chemically equivalent S atoms (S_α and S_β). We expected the S atoms to be closely similar both in intrinsic reactivity and in steric terms given that the disulfide was flanked on both sides by ethylene groups. Now that we have manipulated the regioselectivity of thiol-disulfide interchange by controlling the spatial alignment, it will be of interest to modify innate regioselectivity to probe the interplay between these two factors. This might be done by altering the chemical properties of the substrate disulfide through neighbouring groups, e.g. an electron-withdrawing group or a bulky group might be introduced on either side of the disulfide. A diselenide might also be placed in the substrate, which is expected to react with a nucleophilic thiolate with different kinetics compared to a disulfide (*132*).

4.2.2 Thiol-disulfide interchange in the absence of a pulling force

We have shown unprecedented directionality of motion with the molecular hopper based on thiol-disulfide interchange (see **Chapter 3**). According to the kinetic data we have obtained, we concluded that, at a constant applied potential, the disulfide

bond linking the oligonucleotide with the foothold was forced to react significantly more rapidly with the adjacent cysteine on one side ("downstream cysteine") compared with the other cysteine ("upstream cysteine"). In other words, the pulling force on the DNA effectively shifted the conformational equilibrium of the disulfide. To gain a better understanding of this hopping machinery in the absence of a pulling force, we plan to carry out the following investigations:

1. To probe the behaviour of a hopper molecule without the DNA cargo at the single-molecule level, we will use an uncharged small-molecule disulfide to load a hopper onto the track. It will have a chemical reactivity similar to that of the disulfide in the current hopper-carrier conjugate and will therefore reflect the intrinsic stepping kinetics and directionality in the absence of a pulling force. Previously, it has been shown with an arsenic(III) walker that there is a weak thermodynamic sink towards the trans side of the five-cysteine track, presumably arising from the different reactivities of the five cysteines (*91*).

2. We also plan to study the hopping with the same macromolecular hopper-cargo construct, but under low or no applied transmembrane potential. As discussed in **Chapter 3**, we reasoned that the applied potential (i) set the overall direction by overcoming a large energy barrier to flip the DNA, and (ii) aided in the orientation of the three participating sulfur atoms to favour "forward" reactions to move the charged cargo forward in a field over backsteps.

(i) We have shown that an electrical potential is necessary for flipping the DNA, especially for the switch from cis-to-trans hopping (DNA pulled towards the opening of the barrel—"downwards") to trans-to-cis hopping (DNA pulled towards the vestibule—"upwards"). Intriguingly, two single-stranded DNA (ssDNA) strands must be present inside the pore at the same time during flipping, although a DNA duplex (dsDNA) has

proved too big to translocate through αHL (*133*). It will therefore be of interest to establish a correlation between the voltage applied (e.g. ± 50 mV, ± 75 mV, ± 100 mV) and the time taken to flip the DNA (i.e. the duration of the noisy level after a re-orientating force has been applied, when DNA is presumably migrating through the pore). The voltage-dependence of this flipping step will inform us about the entropic landscape of moving a tethered polymer through a confined tube.

(ii) For the hopping steps, we reasoned that the force contributes to the collinear arrangement of the three participating sulfurs. We plan to validate this hypothesis by conducting a "blind" hopping experiment, where the hopper-DNA is placed on the central foothold of the track before applying zero potential to allow hopping without the pulling force. We will revive the transmembrane potential after a set time (e.g. 5 s, 10 s, 30 s, 1 min) to determine the hopper location. We do not expect equal probability for directional movements, because the polymer cargo is likely to produce other directional biases. For example, we detected unusually large equilibrium constants for hopping steps under negative potentials (e.g. $K_{117\text{-}115}$ > 20 at -150 mV vs $K_{115\text{-}117}$ ~4 at +150 mV), which was attributed to DNA coiling in the cis vestibule and consequently, the disfavouring of backward movement.

4.2.3 Track length

αHL is a well-characterised protein nanopore with an established strategy to generate heteroheptamer containing a single engineered subunit for single-channel recording. It has a β barrel of ~5nm in length (*134*), in which we built the cysteine track containing up to six footholds based on a single β strand. To apply the hopper to characterise long biopolymers, we plan to explore the possibility of building a cysteine track within other β barrel nanopores. One of the candidates is aerolysin, which has a β

barrel of ~10 nm—double in length compared to that of αHL (*130*). Wide-type aerolysin has been used in single-channel recording and exhibited good sensitivity towards the translocating oligonucleotide or polypeptides (*135*, *136*). Therefore, we plan to 1) introduce an octa-aspartate tail to the aerolysin monomer; 2) produce and isolate the heteroheptameric aerolysin containing a single mutant monomer; 3) build a twelve-cysteine track based on aerolysin. In parallel, we plan to explore non-biological surface (e.g. a crystalline surface) to host long tracks.

4.2.4 Computational study on thiol-disulfide interchange under force

While the 'blind' hopping test described above will be useful in unravelling the role of force experimentally, we would also like to model the hopper system in silico to gain further mechanistic insights. Given the complexity of our system where a charged polymer interacts with a protein nanopore under an applied potential, we envisage that coarse-grained molecular dynamics (MD) simulations might reveal which way the disulfide bond is pointing most of the time as a reflection of directionality. Similarly, MD studies can be used to unveil the degree of extension for each substrate (i.e. how stretched a polymer is) (**Chapter 2**) and determine the combination of force and carrier length required for completely regioselective release of the oligonucleotide (>80% regioselectivity has been achieved).

Force has been shown to act as a catalyst in chemistry including thiol-disulfide interchange (*29*, *45*, *51*). In particular, the application of ~200 pN mechanical force induced the prolyl cis-trans isomerisation (ΔE = ~60-80 kJ mol^{-1}) in a model peptide. For a disulfide bond, the shift in C-S-S-C dihedral angle from the preferred 90° to either 0° or 180° significantly increases the energy (ΔE = ~30 kJ mol^{-1} (*7*)), thereby making the disulfide more susceptible to reduction (*137*). It is therefore of interest to probe if a mechanical stress of ~10 pN can overcome the rotational barrier around the S-S bond

axis to accelerate the reaction. In depth computational modelling will be informative to address this question possibly through the hybrid quantum mechanics/molecular mechanics approach. Meanwhile, single-molecule force spectroscopy can be used to validate the computational predictions, by probing the activation of thiol-disulfide interchange with low force (e.g. 1-20 pN).

4.3 Outlooks

In the future, we would like to incorporate the regioselective and site-selective control of thiol-disulfide interchange demonstrated in **Chapter 2** into synthetic tissues based on droplet networks (*83*) as a selective chemical signalling component. Potential downstream responses include the expression of enzymes or the amplification of fluorescence, which can further propagate other forms of signals through the bespoke droplet networks. Coupled with other signalling components (e.g. light and magnetic signals), a multi-layered signalling network can be built within a synthetic tissue, which can be interfaced with biological tissues for functional regulation.

The molecular hopper system described in **Chapter 3** has provided a chemical solution to the stepwise movement of a single-stranded DNA inside the nanopore, which is an alternative to the currently employed biological ratcheting carried out by unwinding enzymes. It is more tolerable to the cargo sequence or even identity compared to enzymes and bears the potential to be transferred to surface. With more intensive research, we hope to develop the hopper into a chemical machine, which can move analytes such as ssDNA, dsDNA, polypeptides, proteins, and polysaccharides across a surface for high-throughput processive characterisation.

References

1. D. Fass, C. Thorpe, Chemistry and enzymology of disulfide cross-linking in proteins. *Chem. Rev.* **118**, 1169–1198 (2018).
2. L. B. Poole, The basics of thiols and cysteines in redox biology and chemistry. *Free Radic. Biol. Med.* **80**, 148–157 (2015).
3. C. C. Winterbourn, M. B. Hampton, Thiol chemistry and specificity in redox signaling. *Free Radic. Biol. Med.* **45**, 549–561 (2008).
4. M. Depuydt, J. Messens, J.-F. Collet, How proteins form disulfide bonds. *Antioxid. Redox Signal.* **15**, 49–66 (2011).
5. N. Nagahara, Catalytic site cysteines of thiol enzyme: sulfurtransferases. *J. Amino Acids*. **2011**, 709404 (2011).
6. P. Nagy, Kinetics and mechanisms of thiol-disulfide exchange covering direct substitution and thiol oxidation-mediated pathways. *Antioxid. Redox Signal.* **18**, 1623–41 (2013).
7. R. Singh, G. M. Whitesides, in *Sulphur-Containing Functional Groups* (John Wiley & Sons, Inc., Chichester, UK, 1993; http://doi.wiley.com/10.1002/9780470034408.ch13), pp. 633–658.
8. C. N. Pace, G. R. Grimsley, J. M. Scholtz, Protein ionizable groups: pK values and their contribution to protein stability and solubility. *J. Biol. Chem.* **284**, 13285–9 (2009).
9. D. A. Dixon, D. J. Zeroka, J. J. Wendoloski, Z. R. Wasserman, The molecular structure of hydrogen disulfide (H2S2) and barriers to internal rotation. *J. Phys. Chem.* **89**, 5334–5336 (1985).
10. R. E. Rosenfield, R. Parthasarathy, J. D. Dunitz, Directional preferences of nonbonded atomic contacts with divalent sulfur. 1. electrophiles and nucleophiles. *J. Am. Chem. Soc.* **99**, 4860–4862 (1977).
11. J. M. Wilson, R. J. Bayer, D. J. Hupe, Structure-reactivity correlations for the thiol-disulfide interchange reaction. *J. Am. Chem. Soc.* **99**, 7922–7926 (1977).
12. G. M. Whitesides, J. Houk, M. A. K. Patterson, Activation parameters for thiolate-disulfide interchange reactions in aqueous solution. *J. Org. Chem.* **48**, 112–115 (1983).

13. J. A. Pappas, Theoretical studies of the reactions of the sulfur-sulfur bond. 1. General heterolytic mechanisms. *J. Am. Chem. Soc.* **99**, 2926–2930 (1977).
14. M. Aida, C. Nagata, An ab initio MO study on the thiol—disulphide exchange reaction. *Chem. Phys. Lett.* **112**, 129–132 (1984).
15. A. Fava, A. Iliceto, E. Camera, Kinetics of the thiol-disulfide exchange. *J. Am. Chem. Soc.* **79**, 833–838 (1957).
16. L. Senatore, E. Ciuffarin, A. Fava, G. Levita, Nucleophilic substitution at sulfur. Effect of nucleophile and leaving group basicity as probe of bond formation and breaking. *J. Am. Chem. Soc.* **95**, 2918–2922 (1973).
17. G. M. Whitesides, J. E. Lilburn, R. P. Szajewski, Rates of thiol-disulfide interchange reactions between mono and dithiols and Ellman's reagent. *J. Org. Chem.* **42**, 332–338 (1977).
18. R. P. Szajewski, G. M. Whitesides, Rate constants and equilibrium constants for thiol-disulfide interchange reactions involving oxidized glutathione. *J. Am. Chem. Soc.* **102**, 2011–2026 (1980).
19. J. A. Bums, G. M. Whitesides, Predicting the Stability of Cyclic Disulfides by Molecular Modeling: "Effective Concentrations" in Thiol-Disulfide Interchange and the Design of Strongly Reducing Dithiols. *J. Am. Chem. Soc.* **112**, 6296–6303 (1990).
20. Z. Shaked, R. P. Szajewski, G. M. Whitesides, Rates of thiol-disulfide interchange reactions involving proteins and kinetic measurements of thiol pKa values. *Biochemistry* **19**, 4156–66 (1980).
21. G. Bulaj, T. Kortemme, D. P. Goldenberg, Ionization-reactivity relationships for cysteine thiols in polypeptides. *Biochemistry* **37**, 8965–8972 (1998).
22. R. Freter, E. R. Pohl, J. M. Wilson, D. J. Hupe, Role of the central thiol in determining rates of the thiol-disulfide interchange reaction. *J. Org. Chem.* **44**, 1771–1774 (1979).
23. D. M. Rothwarf, H. A. Scheraga, Equilibrium and kinetic constants for the thiol-disulfide interchange reaction between glutathione and dithiothreitol. *Proc. Natl. Acad. Sci. U. S. A.* **89**, 7944–8 (1992).
24. P. W. Riddles, R. L. Blakeley, B. Zerner, Ellman's reagent: 5,5′-dithiobis(2-nitrobenzoic acid)-a reexamination. *Anal. Biochem.* **94**, 75–81 (1979).
25. R. D. Bach, O. Dmitrenko, C. Thorpe, Mechanism of thiolate-disulfide

interchange reactions in biochemistry. *J. Org. Chem.* **73**, 12–21 (2008).

26. P. A. Fernandes, M. J. Ramos, Theoretical insights into the mechanism for thiol/disulfide exchange. *Chem. - A Eur. J.* **10**, 257–266 (2004).
27. J. M. Hayes, S. M. Bachrach, Effect of micro and bulk solvation on the mechanism of nucleophilic substitution at sulfur in disulfides. *J. Phys. Chem. A.* **107**, 7952–7961 (2003).
28. R. P. P. Neves, P. A. Fernandes, M. J. Ramos, Mechanistic insights on the reduction of glutathione disulfide by protein disulfide isomerase. *Proc. Natl. Acad. Sci.* **114**, E4724–E4733 (2017).
29. A. P. Wiita, S. R. K. Ainavarapu, H. H. Huang, J. M. Fernandez, Force-dependent chemical kinetics of disulfide bond reduction observed with single-molecule techniques. *Proc. Natl. Acad. Sci.* **103**, 7222–7227 (2006).
30. S. R. Koti Ainavarapu, A. P. Wiita, L. Dougan, E. Uggerud, J. M. Fernandez, Single-molecule force spectroscopy measurements of bond elongation during a bimolecular reaction. *J. Am. Chem. Soc.* **130**, 6479–87 (2008).
31. S. Garcia-Manyes, T. L. Kuo, J. M. Fernández, Contrasting the individual reactive pathways in protein unfolding and disulfide bond reduction observed within a single protein. *J. Am. Chem. Soc.* **133**, 3104–3113 (2011).
32. H. F. Gilbert, Thiol/disulfide exchange equilibria and disulfide bond stability. *Methods Enzymol.* **251**, 8–28 (1995).
33. H. F. Gilbert, Redox control of enzyme activities by thiol/disulfide exchange. *Methods Enzymol.* **107**, 330–51 (1984).
34. M. A. Wouters, S. W. Fan, N. L. Haworth, Disulfides as redox switches: from molecular mechanisms to functional significance. *Antioxid. Redox Signal.* **12**, 53–91 (2010).
35. D. Fass, Disulfide bonding in protein biophysics. *Annu. Rev. Biophys.* **41**, 63–79 (2012).
36. J. Mahatthananchai, A. M. Dumas, J. W. Bode, Catalytic selective synthesis. *Angew. Chem. Int. Ed.* **51**, 10954–10990 (2012).
37. H. J. Davis, R. J. Phipps, Harnessing non-covalent interactions to exert control over regioselectivity and site-selectivity in catalytic reactions. *Chem. Sci.* **8**, 864–877 (2017).
38. K. Inaba *et al.*, Crystal structure of the DsbB-DsbA complex reveals a

mechanism of disulfide bond generation. *Cell* **127**, 789–801 (2006).

39. Y. Li *et al.*, Conformational fluctuations coupled to the thiol-disulfide transfer between thioredoxin and arsenate reductase in Bacillus subtilis. *J. Biol. Chem.* **282**, 11078–83 (2007).
40. G. Dong, P. A. Wearsch, D. R. Peaper, P. Cresswell, K. M. Reinisch, Insights into MHC class I peptide loading from the structure of the tapasin-ERp57 thiol oxidoreductase heterodimer. *Immunity* **30**, 21–32 (2009).
41. K. Yang *et al.*, Crystal structure of the ERp44-peroxiredoxin 4 complex reveals the molecular mechanisms of thiol-mediated protein petention. *Structure* **24**, 1755–1765 (2016).
42. J. Hwang *et al.*, The structural basis for the negative regulation of thioredoxin by thioredoxin-interacting protein. *Nat. Commun.* **5**, 2958 (2014).
43. A. Alon *et al.*, The dynamic disulphide relay of quiescin sulphydryl oxidase. *Nature* **488**, 414–418 (2012).
44. Y. Gat, A. Vardi-Kilshtain, I. Grossman, D. T. Major, D. Fass, Enzyme structure captures four cysteines aligned for disulfide relay. *Protein Sci.* **23**, 1102–1112 (2014).
45. S. Garcia-Manyes, A. E. M. Beedle, Steering chemical reactions with force. *Nat. Rev. Chem.* **1**, 0083 (2017).
46. J. Liang, J. M. Fernández, Mechanochemistry: One bond at a time. *ACS Nano* **3**, 1628–1645 (2009).
47. J. Alegre-Cebollada, P. Kosuri, J. A. Rivas-Pardo, J. M. Fernández, Direct observation of disulfide isomerization in a single protein. *Nat. Chem.* **3**, 882–887 (2011).
48. P. Kosuri *et al.*, Protein folding drives disulfide formation. *Cell* **151**, 794–806 (2012).
49. A. P. Wiita *et al.*, Probing the chemistry of thioredoxin catalysis with force. *Nature* **450**, 124–127 (2007).
50. S. R. K. Ainavarapu *et al.*, Contour length and refolding rate of a small protein controlled by engineered disulfide bonds. *Biophys. J.* **92**, 225–233 (2007).
51. A. E. M. Beedle *et al.*, Forcing the reversibility of a mechanochemical reaction. *Nat. Commun.* **9**, 3155 (2018).
52. R. Perez-Jimenez *et al.*, Diversity of chemical mechanisms in thioredoxin

catalysis revealed by single-molecule force spectroscopy. *Nat. Struct. Mol. Biol.* **16**, 890–896 (2009).

53. T. B. Kahn, J. M. Fernández, R. Perez-Jimenez, Monitoring oxidative folding of a single protein catalyzed by the disulfide oxidoreductase DsbA. *J. Biol. Chem.* **290**, 14518–14527 (2015).
54. D. Giganti, K. Yan, C. L. Badilla, J. M. Fernandez, J. Alegre-Cebollada, Disulfide isomerization reactions in titin immunoglobulin domains enable a mode of protein elasticity. *Nat. Commun.* **9**, 185 (2018).
55. H. Bayley, T. Luchian, S.-H. Shin, M. B. Steffensen, in *Single Molecules and Nanotechnology Springer Series in Biophysics 12*, R. Rigler, H. Vogel, Eds. (2008), pp. 251–277.
56. S. Howorka, H. Bayley, Probing distance and electrical potential within a protein pore with tethered DNA. *Biophys. J.* **83**, 3202–3210 (2002).
57. S.-H. Shin, T. Luchian, S. Cheley, O. Braha, H. Bayley, Kinetics of a reversible covalent-bond-forming reaction observed at the single-molecule level. *Angew. Chem. Int. Ed.* **41**, 3707–3709 (2002).
58. T. Luchian, S.-H. Shin, H. Bayley, Single-molecule covalent chemistry with spatially separated reactants. *Angew. Chem. Int. Ed.* **42**, 3766–71 (2003).
59. S.-H. Shin, H. Bayley, Stepwise growth of a single polymer chain. *J. Am. Chem. Soc.* **127**, 10462–3 (2005).
60. S.-H. Shin, M. B. Steffensen, T. D. W. Claridge, H. Bayley, Formation of a chiral center and pyrimidal inversion at the single-molecule level. *Angew. Chem. Int. Ed.* **46**, 7412–6 (2007).
61. Y. Qing, G. S. Pulcu, N. A. W. Bell, H. Bayley, Bioorthogonal cycloadditions with sub-millisecond intermediates. *Angew. Chem. Int. Ed.* **57**, 1218–1221 (2018).
62. W. J. Ramsay, H. Bayley, Single-molecule determination of the isomers of D-glucose and D-Fructose that bind to boronic acids. *Angew. Chem. Int. Ed.* **57**, 2841–2845 (2018).
63. T. Luchian, S.-H. Shin, H. Bayley, Kinetics of a three-step reaction observed at the single-molecule level. *Angew. Chem. Int. Ed.* **42**, 1926–9 (2003).
64. S. Loudwig, H. Bayley, Photoisomerization of an individual azobenzene molecule in water: an on-off switch triggered by light at a fixed wavelength. *J.*

Am. Chem. Soc. **128**, 12404–5 (2006).

65. A. F. Hammerstein, S.-H. Shin, H. Bayley, Single-molecule kinetics of two-step divalent cation chelation. *Angew. Chem. Int. Ed.* **49**, 5085–5090 (2010).
66. J. Lee, H. Bayley, Semisynthetic protein nanoreactor for single-molecule chemistry. *Proc. Natl. Acad. Sci.* **112**, 13768–13773 (2015).
67. J. Lee *et al.*, Semisynthetic nanoreactor for reversible single-molecule covalent chemistry. *ACS Nano* **10**, 8843–8850 (2016).
68. G. Rousseau, B. Breit, Removable directing groups in organic synthesis and catalysis. *Angew. Chem. Int. Ed.* **50**, 2450–2494 (2011).
69. D. Su, D. A. Jaeger, High regioselectivity in the Diels-Alder reaction of a surfactant 1,3-diene with a surfactant dienophile resulting from a short tether between their functional groups and head groups. *Tetrahedron Lett.* **40**, 7871–7874 (1999).
70. M. Yoshizawa, M. Tamura, M. Fujita, Diels-Alder in aqueous molecular hosts: unusual regioselectivity and efficient catalysis. *Science* **312**, 251–4 (2006).
71. M. T. Huxley *et al.*, Protecting-group-free site-selective reactions in a metal–organic framework reaction vessel. *J. Am. Chem. Soc.* **140**, 6416–6425 (2018).
72. J. bo Wang, G. Li, M. T. Reetz, Enzymatic site-selectivity enabled by structure-guided directed evolution. *Chem. Commun.* **53**, 3916–3928 (2017).
73. S. C. Dodani *et al.*, Discovery of a regioselectivity switch in nitrating P450s guided by molecular dynamics simulations and Markov models. *Nat. Chem.* **8**, 419–425 (2016).
74. K. Kolšek, C. Aponte-Santamaría, F. Gräter, Accessibility explains preferred thiol-disulfide isomerization in a protein domain. *Sci. Rep.* **7**, 9858 (2017).
75. A. Bosco, J. Camunas-Soler, F. Ritort, Elastic properties and secondary structure formation of single-stranded DNA at monovalent and divalent salt conditions. *Nucleic Acids Res.* **42**, 2064–2074 (2014).
76. F. Qin, A. Auerbach, F. Sachs, Estimating single-channel kinetic parameters from idealized patch-clamp data containing missed events. *Biophys. J.* **70**, 264–280 (1996).
77. A. A. Ukuwela, A. I. Bush, A. G. Wedd, Z. Xiao, Glutaredoxins employ parallel monothiol–dithiol mechanisms to catalyze thiol–disulfide exchanges with protein disulfides. *Chem. Sci.* **9**, 1173–1183 (2018).

78. J. Messens *et al.*, How thioredoxin can reduce a buried disulphide bond. *J. Mol. Biol.* **339**, 527–37 (2004).
79. G. Roos *et al.*, How thioredoxin dissociates its mixed disulfide. *PLoS Comput. Biol.* **5** (2009).
80. G. Roos, N. Foloppe, J. Messens, Understanding the pKa of redox cysteines: the key role of hydrogen bonding. *Antioxid. Redox Signal.* **18**, 94–127 (2013).
81. B. Lu, S. Fleming, T. Szalay, J. Golovchenko, Thermal motion of DNA in an MspA pore. *Biophys. J.* **109**, 1439–1445 (2015).
82. V. Noireaux, A. Libchaber, A vesicle bioreactor as a step toward an artificial cell assembly. *Proc. Natl. Acad. Sci.* **101**, 17669–17674 (2004).
83. G. Villar, A. D. Graham, H. Bayley, A tissue-like printed material. *Science* **340**, 48–52 (2013).
84. M. J. Booth, V. R. Schild, A. D. Graham, S. N. Olof, H. Bayley, Light-activated communication in synthetic tissues. *Sci. Adv.* **2** (2016),
85. R. Lentini *et al.*, Integrating artificial with natural cells to translate chemical messages that direct E. coli behaviour. *Nat. Commun.* **5**, 1–6 (2014).
86. M. J. Langton, F. Keymeulen, M. Ciaccia, N. H. Williams, C. A. Hunter, Controlled membrane translocation provides a mechanism for signal transduction and amplification. *Nat. Chem.* **9**, 426–430 (2017).
87. F. G. A. Lister, B. A. F. Le Bailly, S. J. Webb, J. Clayden, Ligand-modulated conformational switching in a fully synthetic membrane-bound receptor. *Nat. Chem.* **9**, 420–425 (2017).
88. M. De Poli *et al.*, Conformational photoswitching of a synthetic peptide foldamer bound within a phospholipid bilayer. *Science*. **352**, 575–580 (2016).
89. V. Janout, S. L. Regen, Bioconjugate-based molecular umbrellas. *Bioconjug. Chem.* **20**, 183–92 (2009).
90. V. Janout, B. Jing, S. L. Regen, Molecular umbrella-assisted transport of thiolated AMP and ATP across phospholipid bilayers. *Bioconjug. Chem.* **13**, 351–356 (2002).
91. G. S. Pulcu, E. Mikhailova, L.-S. Choi, H. Bayley, Continuous observation of the stochastic motion of an individual small-molecule walker. *Nat. Nanotechnol.* **10**, 76–83 (2015).
92. G. Miles, H. Bayley, S. Cheley, Properties of Bacillus cereus hemolysin II: A

heptameric transmembrane pore. *Protein Sci.* **11**, 1813–1824 (2009).

93. M. Montal, P. Mueller, Formation of bimolecular membranes from lipid monolayers and a study of their electrical properties. *Proc. Natl. Acad. Sci. U. S. A.* **69**, 3561–3566 (1972).
94. C. Nicolai, F. Sachs, Solving ion channel kinetics with the QuB software. *Biophys. Rev. Lett.* **08**, 191–211 (2013).
95. W. A. Breyer, B. W. Matthews, A structural basis for processivity. *Protein Sci.* **10**, 1699–1711 (2001).
96. R. D. Vale *et al.*, Direct observation of single kinesin molecules moving along microtubules. *Nature* **380**, 451–453 (1996).
97. R. B. Case, D. W. Pierce, N. Hom-Booher, C. L. Hart, R. D. Vale, The directional preference of kinesin motors is specified by an element outside of the motor catalytic domain. *Cell* **90**, 959–66 (1997).
98. A. Gennerich, R. D. Vale, Walking the walk: how kinesin and dynein coordinate their steps. *Curr. Opin. Cell Biol.* **21**, 59–67 (2009).
99. S. Erbas-Cakmak, D. A. Leigh, C. T. McTernan, A. L. Nussbaumer, Artificial molecular machines. *Chem. Rev.* **115**, 10081–10206 (2015).
100. W. J. Walter, M. P. Koonce, B. Brenner, W. Steffen, Two independent switches regulate cytoplasmic dynein's processivity and directionality. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 5289–93 (2012).
101. R. A. Cross, A. McAinsh, Prime movers: the mechanochemistry of mitotic kinesins. *Nat. Rev. Mol. Cell Biol.* **15**, 257–71 (2014).
102. M. von Delius, E. M. Geertsema, D. A. Leigh, A synthetic small molecule that can walk down a track. *Nat. Chem.* **2**, 96–101 (2010).
103. M. R. Wilson *et al.*, An autonomous chemically fuelled small-molecule motor. *Nature* **534**, 235–40 (2016).
104. S. Erbas-Cakmak *et al.*, Rotary and linear molecular motors driven by pulses of a chemical fuel. *Science* **358**, 340–343 (2017).
105. A. G. Campaña *et al.*, A small molecule that walks non-directionally along a track without external intervention. *Angew. Chem. Int. Ed.* **51**, 5480–5483 (2012).
106. A. G. Campaña, D. A. Leigh, U. Lewandowska, One-dimensional random walk of a synthetic small molecule toward a thermodynamic sink. *J. Am. Chem. Soc.* **135**, 8639–45 (2013).

107. M. J. Barrell, A. G. Campaña, M. von Delius, E. M. Geertsema, D. A. Leigh, Light-driven transport of a molecular walker in either direction along a molecular track. *Angew. Chem. Int. Ed.* **50**, 285–290 (2011).

108. J. E. Beves *et al.*, Toward metal complexes that can directionally walk along tracks: controlled stepping of a molecular biped with a palladium(II) foot. *J. Am. Chem. Soc.* **136**, 2094–100 (2014).

109. Y. He, D. R. Liu, Autonomous multistep organic synthesis in a single isothermal solution mediated by a DNA walker. *Nat. Nanotechnol.* **5**, 778–782 (2010).

110. Y. Tian, Y. He, Y. Chen, P. Yin, C. Mao, A DNAzyme that walks processively and autonomously along a one-dimensional track. *Angew. Chem. Int. Ed.* **44**, 4355–8 (2005).

111. M. You *et al.*, An autonomous and controllable light-driven DNA walking device. *Angew. Chem. Int. Ed.* **51**, 2457–2460 (2012).

112. T.-G. Cha *et al.*, A synthetic DNA motor that transports nanoparticles along carbon nanotubes. *Nat. Nanotechnol.* **9**, 39–43 (2013).

113. K. Yehl *et al.*, High-speed DNA-based rolling motors powered by RNase H. *Nat. Nanotechnol.* **11**, 184–190 (2015).

114. S. F. J. Wickham *et al.*, Direct observation of stepwise movement of a synthetic molecular transporter. *Nat. Nanotechnol.* **6**, 166–169 (2011).

115. H. Gu, J. Chao, S.-J. Xiao, N. C. Seeman, A proximity-based programmable DNA nanoscale assembly line. *Nature* **465**, 202–5 (2010).

116. J.-S. Shin, N. A. Pierce, A synthetic DNA walker for molecular transport. *J. Am. Chem. Soc.* **126**, 10834–5 (2004).

117. B. Ibarra *et al.*, Proofreading dynamics of a processive DNA polymerase. *EMBO J.* **28**, 2794–802 (2009).

118. A. B. C. Deutman *et al.*, Mechanism of threading a polymer through a macrocyclic ring. *Science* **322**, 1668–1671 (2008).

119. N. Koumura, R. W. J. Zijlstra, R. A. van Delden, N. Harada, B. L. Feringa, Light-driven monodirectional molecular rotor. *Nature* **401**, 152–155 (1999).

120. S. Kassem *et al.*, Stereodivergent synthesis with a programmable molecular machine. *Nature* **549**, 374–378 (2017).

121. C. Cheng *et al.*, An artificial molecular pump. *Nat. Nanotechnol.* **10**, 547–553 (2015).

122. C. A. Hunter, H. L. Anderson, What is cooperativity? *Angew. Chem. Int. Ed.* **48**, 7488–7499 (2009).
123. D. A. Keire, E. Strauss, W. Guo, B. Noszal, D. L. Rabenstein, Kinetics and equilibria of thiol/disulfide interchange reactions of selected biological thiols and related molecules with oxidized glutathione. *J. Org. Chem.* **57**, 123–127 (1992).
124. V. M. Krishnamurthy, V. Semetey, P. J. Bracher, N. Shen, G. M. Whitesides, Dependence of effective molarity on linker length for an intramolecular protein−ligand system. *J. Am. Chem. Soc.* **129**, 1312–1320 (2007).
125. J. Nakane, M. Wiggin, A. Marziali, A nanosensor for transmembrane capture and identification of single nucleic Acid molecules. *Biophys. J.* **87**, 615–21 (2004).
126. Y. Wang, K. Tian, L. L. Hunter, B. Ritzo, L.-Q. Gu, Probing molecular pathways for DNA orientational trapping, unzipping and translocation in nanopores by using a tunable overhang sensor. *Nanoscale* **6**, 11372–11379 (2014).
127. J. Venkatraman, G. A. Nagana Gowda, P. Balaram, Design and construction of an open multistranded β-sheet polypeptide stabilized by a disulfide bridge. *J. Am. Chem. Soc.* **124**, 4987–94 (2002).
128. H. Bayley, Nanopore sequencing: from imagination to reality. *Clin. Chem.* **61**, 25–31 (2015).
129. B. Gyarfas *et al.*, Mapping the position of DNA polymerase-bound DNA templates in a nanopore at 5 Å resolution. *ACS Nano* **3**, 1457–66 (2009).
130. I. Iacovache *et al.*, Cryo-EM structure of aerolysin variants reveals a novel protein fold and the pore-formation process. *Nat. Commun.* **7**, 12062 (2016).
131. J. Jiang, B. L. Pentelute, R. J. Collier, Z. H. Zhou, Atomic structure of anthrax protective antigen pore elucidates toxin translocation. *Nature* **521**, 545–549 (2015).
132. H. J. Reich, R. J. Hondal, Why nature chose selenium. *ACS Chem. Biol.* **11**, 821–41 (2016).
133. W. A. Vercoutere *et al.*, Rapid discrimination among individual DNA hairpin molecules at single-nucleotide resolution using an ion channel. *Nat. Biotechnol.* **19**, 248 (2001).
134. L. Song *et al.*, Structure of Staphylococcal alpha -Hemolysin, a Heptameric Transmembrane Pore. *Science* **274**, 1859–1865 (1996).
135. C. Cao *et al.*, Discrimination of oligonucleotides of different lengths with a wild-

type aerolysin nanopore. *Nat. Nanotechnol.* **11**, 713–8 (2016).

136. F. Piguet *et al.*, Identification of single amino acid differences in uniformly charged homopolymeric peptides with aerolysin nanopore. *Nat. Commun.* **9** (2018), doi:10.1038/s41467-018-03418-2.

137. L. Zong *et al.*, Epidithiodiketopiperazines: strain-promoted thiol-mediated cellular uptake at the highest tension. *ACS Cent. Sci.* **3**, 449–453 (2017).

Single-Molecule Studies

International Edition: DOI: 10.1002/anie.201710262
German Edition: DOI: 10.1002/ange.201710262

Bioorthogonal Cycloadditions with Sub-Millisecond Intermediates

*Yujia Qing, Gökçe Su Pulcu, Nicholas A. W. Bell, and Hagan Bayley**

Abstract: *Tetrazine- and sydnone-based click reactions have emerged as important bioconjugation strategies with fast kinetics and N_2 or CO_2 as the only byproduct. Mechanistic studies of these reactions have focused on the initial rate-determining cycloaddition steps. The subsequent N_2 or CO_2 release from the bicyclic intermediates has been approached mainly through computational studies, which have predicted lifetimes of femtoseconds. In the present study, bioorthogonal cycloadditions involving N_2 or CO_2 extrusion have been examined experimentally at the single-molecule level by using a protein nanoreactor. At the resolution of this approach, the reactions appeared to occur in a single step, which places an upper limit on the lifetimes of the intermediates of about 80 μs, which is consistent with the computational work.*

Bioorthogonal click chemistry has emerged as a vital tool in chemical biology. The development of highly selective reactions that proceed in an aqueous environment in near-quantitative yield and involve reagents that are inert towards naturally occurring functional groups has enabled scientists to realise a wide range of applications.[1–3] For example, click chemistry is invaluable for forming conjugates with and between biomacromolecules.

A widely used bioorthogonal process is based on cycloadditions between tetrazines and dienophiles, often referred to as inverse-electron-demand Diels–Alder (iEDDA) reactions.[4] A good match between the $LUMO_{tetrazine}$ and $HOMO_{dienophile}$ endows tetrazine-based iEDDA reactions with fast bimolecular kinetics compared to other click reactions. The rate constants are as high as $10^6\,M^{-1}s^{-1}$ under physiological conditions.[5] The proposed mechanism for tetrazine-based iEDDA reactions iðisplaystyle nvolves an initial cycloaddition to form a bicyclic intermediate, followed by a retro-cycloaddition step in which a molecule of N_2 is extruded to yield a 4,5-dihydropyridazine (Figure 1 a). For cyclooctene dienophiles, slow tautomerisation ensues in protic solvents to yield the 1,4-dihydro isomer.[6,7] Recently, another class of bioorthogonal click reactions, namely between sydnones and strained alkynes, has been developed. This reaction also proceeds through a cycloaddition/retro-cycloaddition pathway, but it is slower (ca. $1\,M^{-1}s^{-1}$).[8,9] In this case, CO_2 is released during the retro-cyclization step (Figure 1 b).

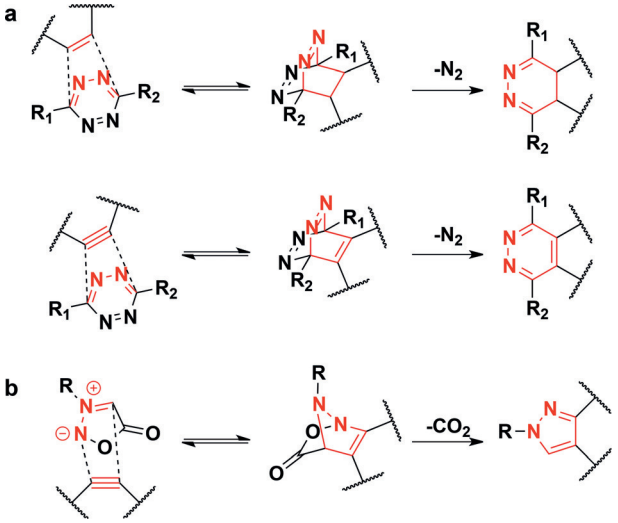


Figure 1. Cycloadditions that release N_2 or CO_2. a) iEDDA reactions between a tetrazine and a strained alkene or alkyne. For the alkene, the subsequent 1,3-prototropic isomerisation step leading to the 1,4-dihydropyridazine is omitted here because the rate is too low for the isomerisation to be detected by the nanoreactor approach.[6,7] b) Cycloaddition between a sydnone and a strained alkyne.

There has been considerable interest in the development of improved bioorthogonal reaction pairs by accelerating the rate-limiting cycloaddition step. Manipulations of the conformational strain of the dienophiles[5,6,10] and the electronic properties of the tetrazines[11,12] have achieved positive outcomes, which have been supported by computational results.[13,14] For example, fusion of the eight-membered ring of *trans*-cyclooctene with a cyclopropane ring accelerated the rate of tetrazine addition by a factor of 27.[6] In contrast, research into the dynamics of N_2 and CO_2 release, which yields the final products, has been largely limited to computational studies,[15] most likely because the intermediates are predicted to be too short-lived to be observed by approaches such as stopped-flow spectroscopy (ca. 1 ms resolution). Indeed, quantum mechanical (QM) and quasi-classical trajectory (QCT) simulations with simplified reactants in the gas phase have suggested that the barrier to N_2 and CO_2 loss is low and that the bicyclic adducts have lifetimes on the order of a C–C bond vibration (ca. 60 fs).[9,15]

Challenged by the computational studies, we set out to seek intermediates for these cycloadditions by our nanoreactor approach, in which individual reactions are carried out within the α-hemolysin (αHL) protein pore in aqueous solution and monitored in real time by ionic current recording.[16] An internal residue within the transmembrane β barrel of the αHL pore is modified with a reactant of interest. Changes in the mass, shape, or polarity of the attached

[*] Y. Qing, Dr. G. S. Pulcu, Dr. N. A. W. Bell, Prof. H. Bayley
Department of Chemistry, University of Oxford
Chemistry Research Laboratory
12 Mansfield Road, Oxford, OX1 3TA (UK)
E-mail: hagan.bayley@chem.ox.ac.uk
Supporting information and the ORCID identification number(s) for the author(s) of this article can be found under: https://doi.org/10.1002/anie.201710262.

 © 2018 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

reactant, caused by bond-making and bond-breaking reactions or isomerisations, alter the ion flow, and cause a change in the recorded current (Figure 2). The single-molecule approach reveals transient intermediates, even in steps that

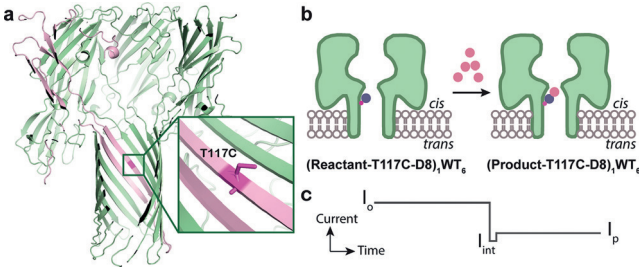


Figure 2. An engineered nanoreactor for single-molecule chemistry. a) Structure of the αHL heteroheptamer used in this work, which contains a single cysteine residue (magenta) at position 117 in one of the seven subunits. b) The heteroheptamer containing one cysteine was modified with a first reactant (purple). The second reactant (pink) was added into the trans solution. Product formation was monitored by ionic current recording. c) Idealised current–time trace showing the expected levels for the open pore (I_o), an intermediate (I_{int}), and the product (I_P).

are not rate-limiting.[17–20] The time resolution of the approach is limited by instrument noise to the microsecond domain.[21,22] With the set-up used in the present work, we expected to have a >90% chance of "seeing" a bicyclic intermediate in a single experiment if its lifetime is 80 μs, and if its formation and disappearance are associated with current steps of 5 pA (see the Supporting Information).

Another advantage of the nanoreactor approach is that the reaction conditions can closely resemble the environment required for bioconjugation. Here, we examine chemistry near a protein surface, in aqueous solution, at neutral pH, and at room temperature. Unlike in silico simulations, the reagents used to perform nanoreactor experiments contain obligatory handles for labelling reactions, enabling direct comparison with the complex scaffolds used in biological applications.

To observe bimolecular click chemistry at the single-molecule level, the αHL pore was first modified with one of the reactants. A single cysteine residue had previously been introduced at position 117 within one of the seven αHL subunits.[17] The side-chain thiol projects into the lumen of the β barrel, and in the present work, reactive groups were introduced into this αHL construct by Michael addition with various maleimide derivatives (Figure 2 and Figure 3a).

For iEDDA reactions involving tetrazines, the "diene" 6-methyltetrazine (mTz) or the dienophile *trans*-cyclooctene (TCO) was attached within the αHL pore to create the nanoreactors (mTz-T117C-D8)$_1$WT$_6$ and (TCO-T117C-D8)$_1$WT$_6$. To examine the reaction of a sydnone with a strained alkyne, the αHL pore was functionalised with a dibenzocyclooctyne (DBCO) to give (DBCO-T117C-D8)$_1$WT$_6$. After a modified pore had been inserted into a lipid bilayer, the open-pore current level was recorded at +50 mV in 2 M KCl, 10 mM MOPS, 10 μM EDTA, pH 7.4. The unitary conductance values were 1.66 ± 0.06 nS for unmodified (T117C-D8)$_1$WT$_6$, 1.19 ± 0.06 nS (28% reduction over the unmodified protein) for (TCO-T117C-D8)$_1$WT$_6$, 1.02 ± 0.10 nS (38% reduction) for (mTz-T117C-D8)$_1$WT$_6$, and 1.46 ± 0.09 nS (12% reduction) for (DBCO-T117C-D8)$_1$WT$_6$.

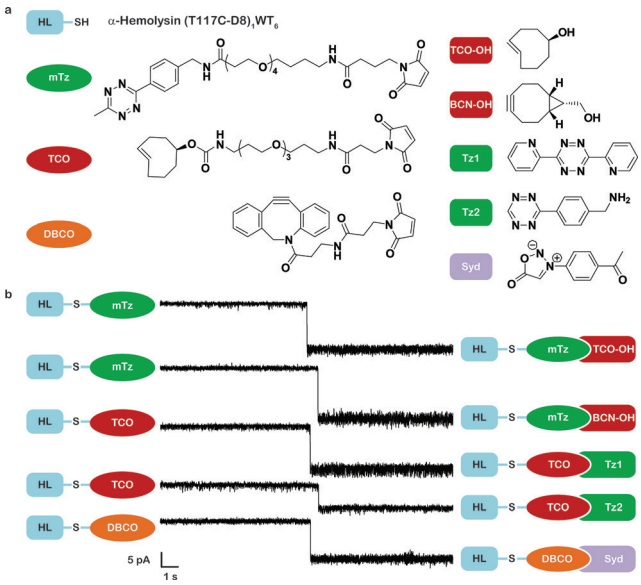


Figure 3. Single-channel recordings for bioorthogonal cycloadditions. a) Structures of the reagents used to modify Cys-117 within the lumen of the nanoreactor (left) and the reactants (right) added to the trans compartment of the bilayer apparatus. b) Current traces showing various cycloaddition reactions. The buffer was 2 M KCl, 10 mM MOPS, pH 7.4, 10 μM EDTA. The temperature was 5 °C for the top trace and room temperature (22 ± 1 °C) for the remaining traces. Note the increases in noise after the cycloaddition reactions.

The addition of a reaction partner to the trans compartment resulted in an additional current blockade (ΔI) after a waiting time (τ) for all five combinations of reactants tested (Figure 3b). Given sufficient individual determinations of τ (e.g., $n = 20$), the mean waiting time $<\tau>$ can be used to determine the bimolecular rate constant, which was carried out for (mTz-T117C-D8)$_1$WT$_6$ and *rel*-(1*R*,4*E*,p*R*)-cyclooct-4-enol (TCO-OH) (see below). The current step was absent when the wild-type αHL pore or unmodified (T117C-D8)$_1$WT$_6$ was used. Neither reversal of the applied potential nor prolongation of the recording time (>2 h) resulted in a reversal of the conductance decrease, indicating that an irreversible reaction had occurred at the reactive site within the pore. No additional current steps were revealed under these conditions, or at pH 6.0 or pH 8.5, suggesting that the reactions either occur in a single step or through a transient (undetected) intermediate.

The reduced current after each reaction was accompanied by increased root-mean-square noise (I_{RMS}; Table 1; see also the Supporting Information, Figure S2), which most likely arises from movement of the reaction product inside the pore. Such noise has been observed previously when a disulfide-bonded polymer was grown within a nanopore,[23] or when molecules such as poly(ethylene glycol)[24] or DNA[25,26] were tethered within a pore. No significant correlation was found between the molecular mass of the reactant and ΔI or ΔI_{RMS},

 © 2018 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

Table 1: Amplitude changes (ΔI) and noise changes (ΔI_{RMS}) after cycloadditions within the αHL nanoreactor.

Modification[a]	Reactant[b]	ΔI [pA]	ΔI_{RMS} [pA]	Repeats
mTz	*rel*-(1R,4*E*,p*R*)-cyclooct-4-enol (a *trans*-cyclooctene, TCO-OH)	11.0 ± 3.0	0.15 ± 0.07	22
	bicyclo[6.1.0]non-4-yn-9-ylmethanol (a strained alkyne, BCN-OH)	12.5 ± 2.0	0.27 ± 0.11	8
TCO	3,6-di-2-pyridyl-1,2,4,5-tetrazine (Tz1)	13.4 ± 1.8	0.36 ± 0.13	6
	3-phenyl-1,2,4,5-tetrazine (Tz2)	9.4 ± 2.0	0.09 ± 0.02	5
DBCO	3-(4-acetyl-phenyl)sydnone (Syd)	12.0 ± 4.0	0.21 ± 0.10	5

[a] The αHL heteroheptamer (Figure 2) was modified to incorporate one of three functional groups: 6-methyltetrazine (mTz), *trans*-cyclooctene (TCO), or dibenzocyclooctyne (DBCO; Figure 3). [b] Reactants were added to the trans compartment. Subsequent cycloaddition led to a reduction in the current and an increase in root-mean-square current noise. Conditions: 2 M KCl, 10 mM MOPS, pH 7.4, 10 μM EDTA at room temperature (22 ± 1 °C).

which depend on additional factors, including the shape, polarity, and flexibility of the modified side chain.[16]

Previously, it had been shown that the release of CO_2 from a carbamic acid in the αHL nanoreactor causes a 5 pA change in current (ΔI) at −50 mV in 2 M KCl.[17] A similar ΔI value might therefore be expected for the N_2 or CO_2 extrusion steps in the cycloadditions examined here. Based on the experimental noise level, the optimal recording conditions for the detection of N_2 or CO_2 extrusion were determined to be a sampling frequency of 25 kHz and a filter frequency of 5 kHz. A four-pole Bessel filter was used with a 10–90 % rise time of 70 μs at 5 kHz. Simulations of individual experiments showed that under these conditions, intermediates with a 5 pA current step and lifetimes of 40 μs, 80 μs, and 120 μs are detected in 23 %, 94 %, and 100 % of the current traces with a 3σ signal-to-noise threshold for detection (see the Supporting Information). For multiple experiments (Table 1), the chance of detecting an intermediate would be still greater.

Reaction pairs involving various substituted dienes and strained dienophiles have been tested (Table 1 and Figure 3). No intermediate species were observed in any case, in multiple experiments ($n > 5$; Table 1 and Figure S3). Therefore, the lifetimes of the bicyclic intermediates must be shorter than the approximately 80 μs detection limit of our experiments. Observation of the reaction between (mTz-T117C-D8)$_1$WT$_6$ and TCO-OH at the lower temperature of 5 °C gave a similar single-step decrease in current. Previously, nanopore interrogation of chemical reactions had revealed intermediates in at least nine cases (Table S4). Events of microsecond duration (5–100 μs) were also revealed in the current traces for interactions between polymers and the αHL pore.[27]

Houk and co-workers have investigated the loss of N_2 during iEDDA reactions of tetrazines by using QM and QCT simulations.[15] Their theoretical results show that a fraction of the bicyclic intermediates with shorter lifetimes than the vibration of the C–C bonds (ca. 60 fs) formed during the cycloadditions is bypassed dynamically, and the reaction is concerted. The authors concluded that the intermediates (fully formed and bypassed) have sub-picosecond scale mean lifetimes, which is about five orders of magnitude below our detection limit. However, the simulations were done in the gas phase, in the absence of collisional relaxation imparted by solvent molecules. The lifetimes of the bicyclic intermediates in the condensed phase remain a question for theoretical chemists to address. Our results provide an experimental upper limit for their values.

For the reactions between (mTz-T117C-D8)$_1$WT$_6$ and TCO-OH, the waiting time before reaction has a mean value $<\tau>$ of 164 ± 20 ms ($n = 20$) at 22 ± 1 °C, yielding a bimolecular rate constant (k) of 14 $M^{-1}s^{-1}$, where $k = 1/(<\tau>[\text{TCO-OH}])$ (Figure S5). Values for similar reactions in the literature range from 80 to 800 $M^{-1}s^{-1}$.[11,28] For the remaining reactions, the rate constants are also estimated to fall within one order of magnitude of the reported values (Table S3). Indeed, it has been shown that a rate constant determined in a nanoreactor can be up to one order of magnitude smaller than the values in bulk solution.[19] This difference has been attributed to the need for the free reactant to find the correct orientation for reaction within a confined space. Our single-molecule approach, with the necessary repeats, has value then for the determination of the overall rate constants for cycloaddition reactions under conditions realistic for bioconjugation.

In conclusion, bioorthogonal cycloaddition chemistry was observed for the first time at the single-molecule level at a protein surface in aqueous solution. The two types of reaction, tetrazine-based iEDDA and sydnone-strained alkyne cycloaddition, were chosen for their similarities in mechanism but distinct extrusion products, N_2 and CO_2. All five cycloaddition reactions examined displayed no intermediates with lifetimes exceeding about 80 μs (94 % certainty in a single experiment; Table 1 and Figure 3). The upper limit on the lifetime of the intermediates that we have determined provides a reference point for future computational studies on bioorthogonal cycloadditions in solution.

Acknowledgements

This work was supported by a European Research Council Advanced Grant. Y.Q. is the recipient of a China Scholarship Council-University of Oxford Scholarship.

 © 2018 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

Conflict of interest

The authors declare no conflict of interest.

Keywords: bioorthogonal chemistry · click chemistry · cycloaddition · single-molecule studies

How to cite: *Angew. Chem. Int. Ed.* **2018**, *57*, 1218–1221
Angew. Chem. **2018**, *130*, 1232–1235

[1] D. M. Patterson, L. A. Nazarova, J. A. Prescher, *ACS Chem. Biol.* **2014**, *9*, 592–605.
[2] J. Li, P. R. Chen, *Nat. Chem. Biol.* **2016**, *12*, 129–137.
[3] F. Liu, Y. Liang, K. N. Houk, *Acc. Chem. Res.* **2017**, *50*, 2297–2308.
[4] M. L. Blackman, M. Royzen, J. M. Fox, *J. Am. Chem. Soc.* **2008**, *130*, 13518–13519.
[5] A. Darko, S. Wallace, O. Dmitrenko, M. M. Machovina, R. A. Mehl, J. W. Chin, J. M. Fox, *Chem. Sci.* **2014**, *5*, 3770–3776.
[6] M. T. Taylor, M. L. Blackman, O. Dmitrenko, J. M. Fox, *J. Am. Chem. Soc.* **2011**, *133*, 9646–9649.
[7] A. Vázquez, R. Dzijak, M. Dračínský, R. Rampmaier, S. J. Siegl, M. Vrabel, *Angew. Chem. Int. Ed.* **2017**, *56*, 1334–1337; *Angew. Chem.* **2017**, *129*, 1354–1357.
[8] S. Wallace, J. W. Chin, *Chem. Sci.* **2014**, *5*, 1742–1744.
[9] M. K. Narayanam, Y. Liang, K. N. Houk, J. M. Murphy, *Chem. Sci.* **2016**, *7*, 1257–1261.
[10] J. Dommerholt, S. Schmidt, R. Temming, L. J. A. Hendriks, F. P. J. T. Rutjes, J. C. M. van Hest, D. J. Lefeber, P. Friedl, F. L. van Delft, *Angew. Chem. Int. Ed.* **2010**, *49*, 9422–9425; *Angew. Chem.* **2010**, *122*, 9612–9615.
[11] M. R. Karver, R. Weissleder, S. A. Hilderbrand, *Bioconjugate Chem.* **2011**, *22*, 2263–2270.
[12] W. Chen, D. Wang, C. Dai, D. Hamelberg, B. Wang, *Chem. Commun.* **2012**, *48*, 1736–1738.
[13] J. A. Wagner, D. Mercadante, I. Nikić, E. A. Lemke, F. Gräter, *Chem. Eur. J.* **2015**, *21*, 12431–12435.
[14] F. Liu, R. S. Paton, S. Kim, Y. Liang, K. N. Houk, *J. Am. Chem. Soc.* **2013**, *135*, 15642–15649.
[15] L. Törk, G. Jiménez-Osés, C. Doubleday, F. Liu, K. N. Houk, *J. Am. Chem. Soc.* **2015**, *137*, 4749–4758.
[16] H. Bayley, T. Luchian, S.-H. Shin, M. B. Steffensen in *Single Molecules and Nanotechnology* (Eds.: R. Rigler, H. Vogel), Springer, Berlin, Heidelberg, **2008**, pp. 251–277.
[17] T. Luchian, S.-H. Shin, H. Bayley, *Angew. Chem. Int. Ed.* **2003**, *42*, 1926–1929; *Angew. Chem.* **2003**, *115*, 1970–1973.
[18] L.-S. Choi, H. Bayley, *Angew. Chem. Int. Ed.* **2012**, *51*, 7972–7976; *Angew. Chem.* **2012**, *124*, 8096–8100.
[19] J. Lee, H. Bayley, *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 13768–13773.
[20] S. Lu, W.-W. Li, D. Rotem, E. Mikhailova, H. Bayley, *Nat. Chem.* **2010**, *2*, 921–928.
[21] M. Mayer, J. K. Kriebel, M. T. Tosteson, G. M. Whitesides, *Biophys. J.* **2003**, *85*, 2684–2695.
[22] G. Shapovalov, H. a. Lester, *J. Gen. Physiol.* **2004**, *124*, 151–161.
[23] S.-H. Shin, H. Bayley, *J. Am. Chem. Soc.* **2005**, *127*, 10462–10463.
[24] S. Howorka, L. Movileanu, X. Lu, M. Magnon, S. Cheley, O. Braha, H. Bayley, *J. Am. Chem. Soc.* **2000**, *122*, 2411–2416.
[25] S. Howorka, L. Movileanu, O. Braha, H. Bayley, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 12996–13001.
[26] S. Howorka, S. Cheley, H. Bayley, *Nat. Biotechnol.* **2001**, *19*, 636–639.
[27] L. Movileanu, H. Bayley, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 10137–10141.
[28] A. Vito et al., *PLoS One* **2016**, *11*, e0167425.

Manuscript received: October 4, 2017
Accepted manuscript online: December 14, 2017
Version of record online: January 5, 2018

 © 2018 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim