Published in final edited form as:

Nat Nanotechnol. 2019 December 01; 14(12): 1135–1142. doi:10.1038/s41565-019-0579-7.

Catalytic site-selective substrate processing within a tubular nanoreactor

Yujia Qing¹, Hiroko Tamagaki-Asahina^{1,2}, Sandra A. Ionescu¹, Mira D. Liu^{1,3,4}, Hagan Bayley^{1,*}

¹Department of Chemistry, University of Oxford, 12 Mansfield Road, Oxford OX1 3TA, UK

³W. M. Keck Science Department, Claremont McKenna College, Claremont, CA 91711, USA

Abstract

Chemists have long sought the ability to modify molecules precisely when presented with several sites of similar reactivity. We reasoned that the confinement of substrates within nanostructures might permit site-selective reactions unachievable in bulk solution, even with sophisticated reagents. In particular, the stretching and alignment of polymers within nanotubes might allow site-specific cleavage or modification. To explore this proposition, macromolecular disulfide substrates were elongated within members of a collection of tubular protein nanoreactors, which contained cysteine residues positioned at different locations along the length of each tube. For each nanoreactor, we defined the reactive location by using a set of polymer substrates (site-selectivity) and which of the two sulfur atoms was attacked (regioselectivity), and found that disulfide interchange occurs with atomic precision. Our strategy has potential for the selective processing of a wide variety of biomacromolecules, and the chemistry and substrates might be generalised yet further by using alternative nanotubes.

The control of selectivity at the nanoscale 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 substrate molecule, and regioselectivity—the ability to distinguish between two (or more) sites within a given functional group^{1,2}. Realising selectivity in synthetic chemistry often requires exploitation

Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use:http://www.nature.com/authors/editorial_policies/license.html#terms

*Correspondence and requests for materials should be addressed to H. Bayley. hagan.bayley@chem.ox.ac.uk.

²Present address: Kyoto Pharmaceutical University, 5 Misasagi-Nakauchi-cho, Yamashina-ku, Kyoto, 607-8414, Japan

⁴Present address: Department of Chemistry, 419 Latimer Hall, University of California, Berkeley, CA 94720-1460

Data Availability statement

The data that support the plots within this paper and other findings of this study are available from the corresponding author upon reasonable request.

Author contributions

YQ and HB conceived the research. YQ performed the experiments. HT contributed to the solid-phase peptide synthesis. SAI and MDL contributed to the electrical measurements. YQ, SAI, and HB wrote the manuscript. All authors discussed the results and commented on the manuscript.

Competing interests

Y.Q. and H.B. have filed patents describing the molecular hopper and applications thereof. The hopper is ratcheted by regioselective thiol-disulfide interchange.

Reprints and permission information is available online at www.nature.com/reprints.

or subversion of the inherent steric and electronic properties of a molecule or functional group². Common strategies include the installation of a reagent-directing group on the substrate, which can be later removed³, or the recruitment of a catalyst that binds to the substrate to favour a particular reaction pathway^{1,2}. Despite significant advances^{1,2}, 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 and therefore hard to generalise².

In enzymes, site-selectivity and regioselectivity are exquisitely executed by the preorganisation of reactants through multiple non-covalent interactions in a confined space—the active site—which can be engineered through directed evolution⁴. For example, a single amino-acid substitution in a cytochrome P450 altered the orientation of its substrate, L-tryptophan, and completely switched the regioselectivity of nitration⁵. To emulate biology, sophisticated chemical catalysts have been developed to bind substrates in defined geometries, notably for site-selective late-stage functionalisation¹. In addition, the preorganisation of reactants, in micelles⁶, or within a supramolecular cage⁷, has been explored to control regioselectivity. However, all these strategies require elaborate effort on a case-by-case basis.

Here, we propose an intuitive strategy to control chemical selectivity by the spatial alignment of reactants within a nanotube (Fig. 1a). Applications of interest include site-selective cleavage and modification of polymeric substrates extended within the tubular nanostructure. Previously, selective chemistry has been observed within carbon nanotubes, including the alignment and orientation of confined monomers to enable polymer formation⁸ and the shielding of reactive sites to direct regioselective aromatic halogenation⁹. Site-selective chemistry on polymers has not been observed. As proof-of-concept, we used a single-molecule approach¹⁰ to examine the site-selectivity and regioselectivity of a model reaction—thiol-disulfide interchange^{11–15}—on macromolecular substrates aligned within tubular nanoreactors.

Experimental approach

To establish whether the nanoscale alignment of reactants within a tube can lead to control over chemical selectivity, we used a system comprising tubular nanoreactors—cysteine mutants of the α -hemolysin (α HL) protein pore (Fig. 1b)—and linear macromolecular substrates with a disulfide bond in the backbone (Fig. 1c). The substrates consisted of biotin—a traptavidin¹⁶-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 segment 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 in both intrinsic reactivity and steric terms.

Four substrates, POC1 to POC4 (Table 1), were tested in which the length of the peptide segment was adjusted two amino acid residues (Gly and Ser) at a time. When a substrate is stretched, each additional dipeptide lowers the disulfide a further ~ 7 Å into the nanotubular

β barrel of the α HL pore (Fig. 1b). A single cysteine was positioned through mutagenesis at seven positions (113, 115, 117, 119, 121, 123 and 143 on one of the seven α HL subunits in each of separate constructs) with the side chain pointing into the lumen of the nanotube where it can take part in thiol-disulfide interchange as a nucleophilic thiolate. Adjacent residues projecting into the pore lumen along a β strand are separated by an average distance of ~ 6.8 Å (C α -C α) and a vertical distance of ~ 5.6 Å, whereas Cys-143 on the antiparallel strand is vertically ~ 2.8 Å lower than Cys-115^{17,18} (Fig. 1b).

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 α HL nanotube from the *cis* compartment under an applied positive potential and was elongated within the electric field. Substrate translocation to the *trans* compartment was prohibited¹⁹ and prolonged immobilization was achieved as the traptavidin-biotin interaction is mechanically stable under the applied force (~ 10 pN)¹⁶. Subsequent alignment-mediated thiol-disulfide interchange with the nanoreactor cysteine side chain resulted in the formation of a covalent adduct within the nanotube (Fig. 1d), which was revealed by a step in the ionic current passing through the pore when the top or bottom part of the substrate dissociated. Dissociation was immediate as no sub-steps were observed. The adduct was then released from the nanoreactor by cleavage with the reducing reagent dithiothreitol (DTT), present in the *trans* compartment, which thereby allowed a new cycle to begin.

Nanoscale alignment controls the reaction between the nanoreactor thiolate and the substrate disulfide

For the moment (see Conclusions), we assume that the system is static, i.e. that there is minimal conformational motion of the protein nanotube and that the substrate remains extended. Previous computational studies supported a collinear transition state for thiol-disulfide interchange, in which adjacent sulfur atoms were separated by ~ 2.5 Å^{20,21}. Assuming free rotation of the nanoreactor cysteine side chains, the vertical distance (Fig. 2a) mapped out by a nucleophilic thiolate is ~ 3 Å (C-S bond length ~ 1.8 Å and C-C-S bond angle $\sim 109.5^\circ$). 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. 2a). 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 unfavourable (Fig. 2a).

The system reported here allows fine tuning of the substrate alignment within a tube at angstrom precision. When the alignment was favourable, the interchange between the substrate disulfide and the nanoreactor thiolate generated a covalent adduct—either an α HL-peptide adduct or an α HL-oligo adduct (discussed later)—which was subsequently released by DTT (Fig. 1d, 3a). When the alignment was unfavourable, the DTT (5 mM) diffusing into the nanopore from the *trans* compartment cleaved the substrate disulfide before interchange could occur; the resulting two fragments exited the pore, the DNA propelled by the applied potential and the peptide segment by diffusion (Fig. 3a). Based on these observations, reactivity was defined—regardless of regioselectivity—as the number of thiol-disulfide interchange reactions between a substrate and a nanoreactor observed under specified

conditions (n_{react}) divided by the total number of substrate threading events (n_{thread}), all of which resulted in interchange or cleavage ($\text{reactivity} = n_{\text{react}}/n_{\text{thread}}$) (Fig. 2b).

Using the aforementioned four substrates (POC1 to POC4) and seven tubular nanoreactors (nanoreactors 113, 115, 117, 119, 121, 123, and 143), we observed a broad range of reactivity from 0% to 100% stemming from the varied alignment (Table 1, Fig. 2). Notably, extension of the peptide segment by two amino acids ($\sim 7 \text{ \AA}$) could dramatically affect 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 the nanoreactor 117 (100% reactivity) (Table 1, Fig. 2). Likewise, shifting the cysteine location by two residues along the β strand ($\sim 5.6 \text{ \AA}$ 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. Therefore, nanoscale alignment can be manipulated by the disulfide position within the substrate or the cysteine location within the nanotube to determine whether or not a reaction proceeds on a substrate.

Nanoscale alignment controls sulfur atom selection within the substrate disulfide

The substrate disulfide connects a neutral segment (the peptide) with a negatively charged polymer (the oligonucleotide) (Fig. 1b). This enabled us to probe whether nanoscale alignment allows substrate processing with atomic precision when the potential reactive sites are within a single bond (i.e. S_{α} and S_{β} in a substrate disulfide). Upon thiol-disulfide interchange within a cysteine nanoreactor, either an α HL-peptide adduct or an α HL-oligo adduct is formed depending on whether the S_{α} or the S_{β} atom of the disulfide bond is attacked. The two adducts showed distinctive current signatures under an applied potential (Fig. 3a), which were assigned based on their behaviour in response to voltage steps (see SI section 1). In contrast to the stable current levels associated with α HL-oligo adducts (S_{β} attacked), rapid switching between three discrete levels was observed for α HL-peptide adducts (S_{α} attacked). We speculate that, in the electric field inside the α HL barrel²², the charged oligonucleotide experiences a constant pulling force that locks its conformation in an extended state, whereas the neutral peptide is less constrained, switching between several conformations. By counting the adducts formed over multiple substrate turnovers, we determined the sulfur regioselectivity for each nanoreactor-substrate combination, and present them as the percentage of interchanges that result in attack on S_{α} or S_{β} as the alignment changes (Fig. 3b, Table 1).

In general, when a nanoreactor cysteine was located within the reactive zone of a substrate disulfide, lowering its position within the nanotube (e.g. from 113 to 115) reduced the chance of forming an α HL-peptide adduct and increased the chance of forming an α HL-oligo adduct. Complete regioselectivity towards S_{β} (α HL-oligo adduct formation) was achieved with POC1 and nanoreactor 115 or 143, POC2 and nanoreactor 119, and POC4 and nanoreactor 121. In these cases, the nucleophilic thiolate could only reach the lower $2.5 \text{ \AA}$

reactive zone of the disulfide (Fig. 2a), solely permitting the transition state leading to an α HL-oligo adduct.

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. 1c) extends 13 Å and that Cys-115 is located ~40 Å from the barrel exit, ~4 phosphodiester bonds of POC1 must be located within the barrel's electric field. Each extension of the peptide segment by two amino acids moves one phosphodiester bond out of the field (the internucleotide distance in stretched single-stranded DNA is 6.9 Å²³). 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} (α HL-peptide adduct formation) was achieved with POC4 and nanoreactor 113. We envisage that complete S_{α} regioselectivity might be achieved by using a longer peptide segment and a longer pore.

Nanoscale alignment controls the kinetics of thiol-disulfide interchange

After a substrate threads into a nanoreactor, one of three possible reactions occurs: generation of an α HL-peptide adduct, generation of an α HL-oligo adduct, or cleavage by DTT. At the single molecule level, rate constants are related to the mean dwell times of states¹⁰. For the formation of an adduct within a molecular complex, the rate constant— k_{α} for the formation of the α HL-peptide adduct and k_{β} for the formation of the α HL-oligo adduct—is the reciprocal of the corresponding mean dwell time (were only one of the reactions to occur) ($\langle \tau_{\alpha} \rangle$ or $\langle \tau_{\beta} \rangle$, e.g. $\langle \tau_{\alpha} \rangle = 1/k_{\alpha}$). For the bimolecular DTT cleavage at a fixed DTT concentration (e.g. 5 mM), $\langle \tau_{\text{DTT}} \rangle = 1/k'_{\text{DTT}} = 1/k_{\text{DTT}} [\text{RS}^-]_{\text{DTT}}$, where k'_{DTT} is the pseudo-first order rate constant incorporating $[\text{RS}^-]_{\text{DTT}}$, the concentration of the DTT monothiolate (at pH 8.5, the concentration of doubly ionised DTT is negligible²⁴).

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

$$\langle \tau \rangle = \frac{1}{k_{\alpha} + k_{\beta} + k'_{\text{DTT}}} \quad (1)$$

$$\langle \tau \rangle = \frac{\langle \tau_{\alpha} \rangle \langle \tau_{\beta} \rangle \langle \tau_{\text{DTT}} \rangle}{\langle \tau_{\alpha} \rangle \langle \tau_{\beta} \rangle + \langle \tau_{\beta} \rangle \langle \tau_{\text{DTT}} \rangle + \langle \tau_{\alpha} \rangle \langle \tau_{\text{DTT}} \rangle} \quad (2)$$

Depending on the alignment, one of the three pathways may dominate, e.g. $\langle \tau_{\alpha} \rangle, \langle \tau_{\text{DTT}} \rangle \gg \langle \tau_{\beta} \rangle$, in which case $\langle \tau \rangle = \langle \tau_{\beta} \rangle$. In cases where more than one pathway compete, dwell-time analysis²⁵ was performed to derive the independent rate constants ($k_{\alpha}, k_{\beta}, k'_{\text{DTT}}$) (Table 1).

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 barrel²⁶. Nevertheless, when comparing different substrates reacting with the same nanoreactor (i.e. the cysteine pK_a is fixed) at a fixed pH (e.g. pH 8.5), k_{α} and k_{β} directly correlate with the thiolate

alignment with respect to S_α and S_β , and the regioselectivity is given by $k_\beta/k_\alpha = n_\beta/n_\alpha$, where n_β is the number of α HL-oligo adducts formed in a time interval and n_α the number of α HL-peptide adducts formed in the same interval. The spatial relationships between the substrates and a particular nanoreactor can be inferred from the derived rate constants (Table 1). For each nanoreactor, optimal alignment with the substrate S_β atom was required to obtain the highest rate of α HL-oligo adduct formation (k_β). The highest k_β values were between nanoreactor 113 and POC1, nanoreactor 115 and POC3, nanoreactor 117 and POC3 and nanoreactor 119 with POC4 (Table 1). For example, a ~ 300 -fold range in k_β values was seen with nanoreactor 117, which gave $5.3 \times 10^{-4} \text{ s}^{-1}$ with POC1 and 0.17 s^{-1} with POC3. This shifts the overall outcome from domination by DTT cleavage with POC1 ($k'_{\text{DTT}} = 9.4 \times 10^{-3} \text{ s}^{-1}$) to domination by thiol-disulfide interchange at S_β with POC3, showcasing the influence of alignment on reaction kinetics and site-selectivity. As expected, DTT cleaved the various substrates at similar rates, which were independent of the cysteine position in the nanotube (Table 1). Hence the reactivity (Fig. 2b) and the rates of adduct formation (k_α and k_β , Table 1) are correlated. Further, the kinetic data are consistent with alignments obtained from modelling studies in which the substrate chain was extended (Fig. 2c).

Site-selective and regioselective tandem thiol-disulfide interchanges

To recapitulate the tandem thiol-disulfide interchanges frequently seen with protein dithiols^{27,28}, we constructed two double-cysteine nanoreactors based on nanoreactor 115 with an additional cysteine on either the antiparallel strand (115/143) or on the same strand (115/117) within a single α HL subunit. Using POC1—which reacted with the three single-cysteine nanoreactors (115, 143, and 117) with complete S_β regioselectivity—we demonstrated that consecutive thiol-disulfide interchanges occur within the double-cysteine nanotubes.

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 α HL-oligo adduct. With nanoreactor 115/117, α HL-oligo-adducts at positions 115 and 117 can be differentiated by their residual currents ($\text{Ires}\%_{(117)} - \text{Ires}\%_{(115)} = +2.0\% \pm 0.5\%$, $n = 3$) (Fig. 4a), which is in accord with the results obtained with single-cysteine nanoreactors ($\text{Ires}\%_{(115)} = 45\% \pm 1\%$, $n = 9$; $\text{Ires}\%_{(117)} = 47\% \pm 1\%$, $n = 12$). With nanoreactor 115/143, α HL-oligo adducts at positions 115 and 143 can also be differentiated ($\text{Ires}\%_{(115)} - \text{Ires}\%_{(143)} = +0.68\% \pm 0.03\%$, $n = 3$) (Fig. 4b and Supplementary Fig. 2), which again is consistent with results with single-cysteine nanoreactors ($\text{Ires}\%_{(115)} = 45\% \pm 1\%$, $n = 9$; $\text{Ires}\%_{(143)} = 46\% \pm 1\%$, $n = 6$). With nanoreactor 115/117, we found that the substrate disulfide reacted site-selectively (100%) with Cys-115 to form the α HL-oligo adduct ($n = 30$), which is in accord with the kinetics of the single-cysteine nanoreactors: $k_{\beta(115)} \gg k_{\beta(117)}$. In contrast, we observed 81% reaction with Cys-115 and 19% reaction with Cys-143 with nanoreactor 115/143, also consistent with the kinetics of the single-cysteine nanoreactors (Table 1, Fig. 4b, Supplementary Fig. 2).

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 (S_γ) or the sulfur contributed by the oligonucleotide

(S_{β}) (Fig. 4). Interestingly, 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. 4a). When the free cysteine was on the antiparallel strand (115/143), we observed thiolate attack on both S_{β} and S_{γ} (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. 4b). The subsequent addition of DTT regenerated free cysteines through a transient α HL-DTT intermediate²⁹ (Fig. 4b, see SI section 3). Among the α HL-oligo adducts formed on Cys-115, 85% were immediately released through intramolecular S_{γ} attack by Cys-143 without transfer.

The regioselectivities of the second thiol-disulfide interchanges were attributed to the C_{α} - C_{α} distances between the two cysteines involved (Fig. 4 and Supplementary Fig. 2). For nanoreactor 115/117, in which C_{α} - C_{α} = 6.8 Å, S_{β} attack occurred exclusively (Fig. 4). For nanoreactor 115/143, in which C_{α} - C_{α} = 4.6 Å, S_{γ} attack predominated and was more favourable when the adduct was on Cys-115 (Fig. 4).

Conclusions

In the present work, we confined and extended a polymer substrate within a nanotube to promote site-selective and regioselective chemistry. Thiol-disulfide interchange, monitored at the single-molecule level, revealed whether nanoscale alignment determined that a substrate undergoes reaction and if so at which site. Notably, selection between two chemically equivalent sulfur atoms within a single disulfide bond (S_{α} and S_{β}) was achieved, which does not occur in bulk solution. By manipulating the relative positions of the substrate disulfide and the thiolate in the nanotube, we steered substrate processing between complete S_{β} regioselectivity and >80% S_{α} regioselectivity. With double-cysteine nanoreactors, tandem thiol-disulfide interchanges exhibited both remarkable site-selectivity and regioselectivity, which occurs in biological systems²⁷ but is again unprecedented in bulk solution. The ability of an α HL-oligo adduct to move back and forth between two Cys residues has been further developed to construct a molecular "hopper" that can advance up and down a 5- or 6-cysteine track³⁰.

To aid description, we have so far assumed that our reaction system is static. However, we expect significant molecular motion of the components, especially the confined polymer substrate³¹. In a static system, site-selectivity and regioselectivity would be achieved by accurate alignment and consequently, reactions would occur at rapid rates. However, just as in host-guest systems, perfect organisation can be difficult to achieve. Further, a degree of conformational flexibility can be advantageous. Indeed, we found that complete S_{β} regioselectivity was always achieved among substrates with the lowest k_{β} values (Table 1, Fig. 3b,c). In these cases, it seems likely that the reactive thiol in the nanotube is positioned on average below the substrate disulfide (i.e. towards the *trans* side). Then, rare encounters in a collinear arrangement resulting from conformational flexibility result in S_{β} regioselectivity, while S_{α} is inaccessible (Fig. 3c). We therefore have also revealed dynamic details underlying this nanoscale system, which will be informative for future elaboration.

In future work, the steric and electronic properties of the functional groups flanking the disulfide in the substrate might be manipulated during synthesis to produce innate regioselectivity, and it would be interesting to probe the competition with alignment in such biased cases. In addition, the current system might be developed into an array of membrane-embedded protein nanotubes for parallel substrate processing, which has been adopted in nanopore sequencing devices³². The fact that the system is catalytic and permits multiple turnovers (Fig. 1d) will further boost the substrate throughput.

The strategy established here might be applied to a broad spectrum of substrates and chemistries. Given their biomedical relevance, biopolymers (e.g. polypeptides, oligonucleotides, fatty acids, polysaccharides, linear polyketides) are important candidates for site-selective chemistry. The diverse functionalities within biopolymers present numerous options for processing chemistries. For example, immediate applications of the protein nanopore-based system could be centred around biocompatible chemistries with thiols, amines, and thiophosphates (as in Cys and Lys protein side chains, and phosphorothioate oligonucleotides). By using nanotube alignment, a functional group, for example a protecting group or a fluorescent label, can be added to or removed from a specific site in a biopolymer. Sequential cleavage and ligation, which occurs in RNA and protein splicing, might also be realised.

Recent advances have generated ideal tubular nano-scaffolds, such as metal-organic frameworks, covalent-organic frameworks and carbon nanotubes, within which molecules can be organised. For example, a symmetrical dialkyne was recently converted to a monotriazole, by the reaction with azide groups spaced within a metal-organic framework³³. We envision that the strategy reported here can be adapted to these synthetic nano-systems to allow rapid site-selective or regioselective turnover of substrates confined within them.

Methods

Single-molecule characterisation of substrate processing within a nanotube

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 $^{\circ}\text{C}$) for 15 min to form the traptavidin-tagged substrate. A planar bilayer was first formed across the aperture in a recording chamber by using the method established by Montal and Mueller (see SI section 9). The *cis* compartment of the recording chamber (see SI section 9) containing αHL nanoreactor heptamers was stirred until a single pore inserted into the bilayer, as seen on the current readout. 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 the same buffer supplemented with 5 mM DTT. To drive the macromolecular substrate into the nanopore, a potential of +150 mV was applied to the *trans* side of the bilayer. The threaded substrate reduced the current to an $I_{\text{res}\%}$ of $\sim 33\%$ (see SI section 4 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 from the nanoreactor (see SI section 4 for $I_{\text{res}\%}$ for each substrate-nanoreactor combination).

Subsequent reaction with DTT released the adduct and returned the nanoreactor to its initial state ready for another round of thiol-disulfide interchange.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank Lapatrada Taemaitree for her help with LC-MS characterization. We also thank Dr. Michael Booth for useful discussions and Professor Mark Howarth for providing a sample of traptavidin for preliminary tests. This research was supported by a European Research Council Advanced Grant (COSIMO). Y.Q. was supported by a China Scholarship Council-University of Oxford Scholarship. S.A.I. was funded by an Amelia Jackson Studentship, Exeter College, Oxford.

References

1. Davis HJ, Phipps RJ. Harnessing non-covalent interactions to exert control over regioselectivity and site-selectivity in catalytic reactions. *Chem Sci*. 2017; 8:864–877. [PubMed: 28572898]
2. Mahatthananchai J, Dumas AM, Bode JW. Catalytic selective synthesis. *Angew Chem Int Ed*. 2012; 51:10954–10990.
3. Rousseau G, Breit B. Removable directing groups in organic synthesis and catalysis. *Angew Chem Int Ed*. 2011; 50:2450–2494.
4. Wang J, Li G, Reetz MT. Enzymatic site-selectivity enabled by structure-guided directed evolution. *Chem Commun*. 2017; 53:3916–3928.
5. Dodani SC, et al. Discovery of a regioselectivity switch in nitrating P450s guided by molecular dynamics simulations and Markov models. *Nat Chem*. 2016; 8:419–425. [PubMed: 27102675]
6. Su D, Jaeger DA. 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*. 1999; 40:7871–7874.
7. Yoshizawa M, Tamura M, Fujita M. Diels-Alder in aqueous molecular hosts: unusual regioselectivity and efficient catalysis. *Science*. 2006; 312:251–4. [PubMed: 16614218]
8. Nakanishi Y, et al. Template synthesis of linear-chain nanodiamonds inside carbon nanotubes from bridgehead-halogenated diamantane precursors. *Angew Chem Int Ed*. 2015; 54:10802–10806.
9. Miners SA, Rance GA, Khlobystov AN. Regioselective control of aromatic halogenation reactions in carbon nanotube nanoreactors. *Chem Commun*. 2013; 49:5586.
10. Bayley, H, Luchian, T, Shin, S-H, Steffensen, MB. Single-molecule covalent chemistry in a protein nanoreactor. *Single Molecules and Nanotechnology Springer Series in Biophysics* 12. Rigler, R, Vogel, H, editors. 2008. 251–277.
11. Alegre-Cebollada J, Kosuri P, Rivas-Pardo JA, Fernández JM. Direct observation of disulfide isomerization in a single protein. *Nat Chem*. 2011; 3:882–887. [PubMed: 22024885]
12. Beedle AEM, et al. Forcing the reversibility of a mechanochemical reaction. *Nat Commun*. 2018; 9:3155. [PubMed: 30089863]
13. Beedle AEM, Mora M, Lynham S, Stirnemann G, Garcia-Manyes S. Tailoring protein nanomechanics with chemical reactivity. *Nat Commun*. 2017; 8
14. Beedle AEM, Lynham S, Garcia-Manyes S. Protein S-sulfonylation is a fleeting molecular switch that regulates non-enzymatic oxidative folding. *Nat Commun*. 2016; 7:1–10.
15. Garcia-Manyes S, Beedle AEM. Steering chemical reactions with force. *Nat Rev Chem*. 2017; 1:0083.
16. Chivers CE, et al. A streptavidin variant with slower biotin dissociation and increased mechanostability. *Nat Methods*. 2010; 7:391–393. [PubMed: 20383133]
17. Murzin AG, Lesk AM, Chothia C. Principles determining the structure of beta-sheet barrels in proteins. *J Mol Biol*. 1994; 236:1382–400. [PubMed: 8126727]

18. Song L, et al. Structure of staphylococcal α -hemolysin, a heptameric transmembrane pore. *Science*. 1996; 274:1859–66. [PubMed: 8943190]
19. Stoddart D, Heron AJ, Mikhailova E, Maglia G, Bayley H. Single-nucleotide discrimination in immobilized DNA oligonucleotides with a biological nanopore. *Proc Natl Acad Sci U S A*. 2009; 106:7702–7707. [PubMed: 19380741]
20. Fernandes PA, Ramos MJ. Theoretical insights into the mechanism for thiol/disulfide exchange. *Chem Eur J*. 2004; 10:257–266. [PubMed: 14695571]
21. Bach RD, Dmitrenko O, Thorpe C. Mechanism of thiolate-disulfide interchange reactions in biochemistry. *J Org Chem*. 2008; 73:12–21. [PubMed: 18052192]
22. Howorka S, Bayley H. Probing distance and electrical potential within a protein pore with tethered DNA. *Biophys J*. 2002; 83:3202–3210. [PubMed: 12496089]
23. Bosco A, Camunas-Soler J, Ritort F. Elastic properties and secondary structure formation of single-stranded DNA at monovalent and divalent salt conditions. *Nucleic Acids Res*. 2014; 42:2064–2074. [PubMed: 24225314]
24. Whitesides GM, Lilburn JE, Szajewski RP. Rates of thiol-disulfide interchange reactions between mono and dithiols and Ellman's reagent. *J Org Chem*. 1977; 42:332–338.
25. Qin F, Auerbach A, Sachs F. Estimating single-channel kinetic parameters from idealized patch-clamp data containing missed events. *Biophys J*. 1996; 70:264–280. [PubMed: 8770203]
26. Pace CN, Grimsley GR, Scholtz JM. Protein ionizable groups: pK values and their contribution to protein stability and solubility. *J Biol Chem*. 2009; 284:13285–9. [PubMed: 19164280]
27. Depuydt M, Messens J, Collet J-F. How proteins form disulfide bonds. *Antioxid Redox Signal*. 2011; 15:49–66. [PubMed: 20849374]
28. Ukuwela AA, Bush AI, Wedd AG, Xiao Z. Glutaredoxins employ parallel monothiol–dithiol mechanisms to catalyze thiol–disulfide exchanges with protein disulfides. *Chem Sci*. 2018; 9:1173–1183. [PubMed: 29675162]
29. Luchian T, Shin S-H, Bayley H. Single-molecule covalent chemistry with spatially separated reactants. *Angew Chem Int Ed*. 2003; 42:3766–71.
30. Qing Y, Ionescu SA, Pulcu GS, Bayley H. Directional control of a processive molecular hopper. *Science*. 2018; 361:908–912. [PubMed: 30166488]
31. Lu B, Fleming S, Szalay T, Golovchenko J. Thermal motion of DNA in an MspA pore. *Biophys J*. 2015; 109:1439–1445. [PubMed: 26445444]
32. Bayley H. Nanopore sequencing: from imagination to reality. *Clin Chem*. 2015; 61:25–31. [PubMed: 25477535]
33. Huxley MT, et al. Protecting-group-free site-selective reactions in a metal–organic framework reaction vessel. *J Am Chem Soc*. 2018; 140:6416–6425. [PubMed: 29699391]

Fig. 1. Substrate processing within a protein nanotube.

a, In the present study, a strategy enabling site-selective cleavage or modification of a polymer, which does not occur in bulk, is demonstrated by confining and aligning the polymer substrate within a nanotube. **b**, The transmembrane β barrel is ~ 2 nm in internal diameter and ~ 5 nm in height. The lumen of the pore is lined by generally hydrophilic side chains. Details can be found in the reported high-resolution structure¹⁸. Cysteines (pink) on one of the seven α -hemolysin (α HL) subunits provide nucleophilic thiolates (yellow) at defined sites within the transmembrane β barrel. Six single-cysteine mutants (113, 115, 117, 119, 121, 143) and two double-cysteine mutants (115/117 and 115/143) were used. **c**, Substrates consisted of a biotin at one terminus followed by a peptide-oligonucleotide junction formed by click chemistry. The cleavable disulfide bond (purple and pink dots) was installed between the click product and the oligo. The peptide length was manipulated by

adding two amino acids (glycine or serine, $n = 2$) at a time, following the PEG unit: POC1 ($n = 0$); POC2 ($n = 2$), POC3 ($n = 4$), POC4 ($n = 6$). **d**, A molecule of traptavidin-capped substrate threads into the nanopore (i) from the *cis* side under an applied potential (+150 mV, *trans*), and translocation is arrested by the traptavidin stopper (ii). Favourable alignment enables thiol-disulfide interchange between the cysteine thiolate and the substrate disulfide, generating either an α HL-peptide adduct (top, iv) or an α HL-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.

Fig. 2. Nanoscale 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 $\sim 180^\circ$. 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). Alignment of 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 = $n_{\text{interchange}}/n_{\text{thread}}$) 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 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.

Fig. 3. Regioselectivity of single-molecule thiol-disulfide interchange.

a. Identification of turnover intermediates from current signatures. Threading of POC1 into nanoreactor 113 from the *cis* compartment (i) 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 pattern (iii, iv) corresponding to the formation of two different covalent adducts within the nanopore (see SI section 1): the α HL-peptide adduct offered three sub-conductance states (iii: purple lines: $I_{res}\% = 43\%, 48\%, 51\%$); the α HL-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 μ M EDTA, 5 mM DTT (*trans*), pH 8.5, 20 ± 1 °C. **b**, The sulfur regioselectivity of thiol-disulfide interchange between substrates POC1 (red), POC2 (blue), POC3 (yellow), and POC4 (green) and nanoreactors 113, 115, 143, 117, 119, and 121, expressed as a percentage of the total number of adducts formed (reaction at S_{α} and S_{β}). **c**, The relative positions of the substrates and the cysteines within single-cysteine nanoreactors inferred from the regioselectivity data. The area of each solid circle is scaled to the regioselectivity ratio (S_{β}/S_{α}) for the corresponding combination of substrate and nanoreactor. Empty circles indicate a complete S_{β} regioselectivity. The disulfides are placed where S_{β}/S_{α} regioselectivity = 1, as extrapolated from the regioselectivity data. Accessible positions of the disulfide are indicated with dashed lines based on the reactivity data (Fig. 2c).

Fig. 4. Tandem thiol-disulfide interchanges with double-cysteine nanoreactors.

a, With nanoreactor 115/117, POC1 reacted site-selectively and regioselectively with cysteine 115 to form an α HL-oligo adduct (iv). 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 α HL-DTT adduct (blue arrow). In the presence of DTT, the average number of transfers per cycle before adduct release by DTT was 11 ($n = 40$ cycles) and up to 29 successive transfers were

recorded in a single cycle. **b**, With nanoreactor 115/143, POC1 reacted with 81% site-selectivity and 100% regioselectivity at cysteine 115 to form an α HL-oligo adduct (iv). 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 (ii, substrate within pore) until DTT regenerated the two free cysteines through an intermediate α HL-DTT adduct (blue arrow). Attack on S_β by Cys-143 was seen occasionally, resulting in adduct transfer (see SI section 2). The average number of transfers before intramolecular release depended on which site an α HL-oligo adduct initially formed: 0.54 transfers per cycle ($n = 48$ cycles) for the α HL-oligo adducts initially formed on Cys-115 and 1.36 ($n = 11$ cycles) for those initially formed on Cys-143. The difference of ~ 1 implies that release was exclusively from Cys-115, which was indeed the case (Supplementary Fig. 2). Traces were filtered at 200 Hz. Insets showing α HL-DTT adducts were filtered at 1000 Hz. Conditions: 2 M KCl, 20 mM HEPBS, 20 μ M EDTA, 5 mM DTT (*trans*), pH 8.5, 20 ± 1 °C.

Table 1
Reactivity, selectivity and kinetics of thiol-disulfide interchange reactions

Substrate ^a	Peptide segment ^b	Nano-reactor	Reactivity	Regioselectivity ^c		$k_{\beta} / \text{s}^{-1d}$	$k_{\alpha} / \text{s}^{-1d}$	$k_{\text{DTT}} / \text{s}^{-1d}$
				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	/
		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 All substrates contained an oligoadenosine 40mer (A40).

^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 α HL-oligo adduct (S_{β}) or an α HL-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 software.

^e /, not measurable.

^f DTT cleavage occurred < 5 times, and hence was disregarded in the kinetic model.