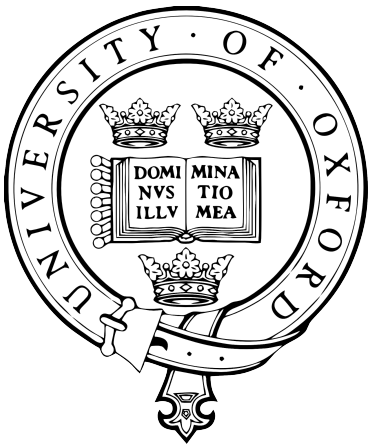
Developing single-molecule imaging and spectroscopy techniques to probe molecular properties and interactions



Timothy J D Bennett

Merton College

University of Oxford

This thesis is submitted for the degree of

Doctor of Philosophy

Physical and Theoretical Chemistry Department

Trinity Term 2024

Acknowledgements

I would like to thank Prof. Krishnan for the opportunity to work with such an exciting technology. You have been an inspirational scientist to me and a constant source of encouragement and guidance, thank you for the trust and support throughout.

It takes a team to make a research lab, and much of the work presented in this thesis has absolutely been a team effort, so I have to thank all my teammates in the group, without whom these past years would not have been the same. You've always been on hand to provide me with laughs, advice and interesting discussions, during coffee breaks and volleyball trainings alike.

I could not have asked for a better bunch of friends to share my Oxford experience with. Tom, Freddie and Archie, you've been the best housemates, and I could always rely on you to distract me from whatever went wrong in the lab that week, whether through a squash match, a quick Spirit, or a cheeky line-out.

Finally, Ana, you've been my rock throughout, constantly providing me with perspective. I'm so glad we could go on our DPhil journeys together.

Table of contents

1 Introduction **3**

1.1 Why study single molecules? . . . 3

1.2 Single-molecule methods . . . 4

1.3 Effective electrical charge in solution . . . 6

1.4 Escape-time electrometry (ET*e*) . . . 9

1.5 Thesis outline . . . 16

1.6 List of publications . . . 17

2 Ensemble ET*e* **19**

2.1 Measurements of 3D molecular conformation . . . 19

2.2 Measurements of post-translational modifications (PTMs) . . . 29

3 Single-molecule ETe **37**

3.1 Trajectory-based ET*e* (TrET*e*) . . . 37

3.2 Flow-based ET*e* (FlET*e*) . . . 43

4 Escape time size-and-shape spectrometry (ETsy) **53**

4.1 Escape-time measurements under physiological conditions . . . 53

5 Conclusion **85**

Appendix A Device fabrication **89**

Appendix B Widefield epifluorescence microscopy **93**

Appendix C Escape-time movie analysis **99**

Appendix D Brownian dynamics (BD) simulations of escape-time experiments **103**

D.1 BD simulations of ET*e* experiments . . . 103

D.2 BD simulations of FlET*e* experiments 105

D.3 BD simulations of ETsy experiments 106

Appendix E Experimental error in determining q_{eff} **109**

Appendix F The effect of SNR and photo-blinking on escape-time measurements **111**

Appendix G Supporting information for "Far-field electrostatic signatures of macromolecular 3D conformation" **115**

Appendix H Supplementary materials for "Single-molecule size-and-shape spectrometry on a chip" **155**

References **217**

Abstract

This thesis outlines recent advancements to methods of single-molecule measurement in solution, building upon the existing escape-time electrometry (ETe) technology. ETe uses a field-free fluid trap for high resolution measurements of molecular effective charge in the solution phase, exemplified here with measurements of 3D macromolecular conformation, and the detection of post-translational modifications in proteins. Such measurements showcase the capability of ETe to address poignant biological applications, such as disease state detection. Two single-molecule realisations of the ETe technology are presented, which facilitate quantification of heterogeneity in a sample through single-particle tracking, and translocation times across a trapping nanostructure respectively. When combined with the sensitivity of ETe, these single-molecule methods will enable future measurements of currently challenging systems, such as the conformation states of ribonucleic acids. Finally, ETe measurements of size, conformation and reaction kinetics in physiological conditions are presented. These measurements on molecules over three orders of magnitude in molecular weight, and at over five orders of magnitude in concentration, lay the foundation for a multitude of applications from structural and cell biology to diagnostics and physical chemistry.

Chapter 1

Introduction

1.1 Why study single molecules?

Single-molecule measurements have been shown to provide valuable insight into processes across a wide range of fields, from photophysics to surface chemistry to cell biology[1]. The ability to interrogate individual molecules rather than relying upon an ensemble average measurement enables at least three major advantages. Firstly, quantification of the heterogeneity of a sample is possible; the properties of underlying distributions of states can be measured, including the identification of rare states which ensemble methods may struggle to detect[2]. Fluctuations in these states can also be probed, whether initiated by the addition of external stimuli or due to stochastic inter-conversion between the inherent states of the sample. One common example of this is the complex molecular machines of molecular motors whose kinetics are hard to synchronise in large ensembles, and so measurements of their operation require single-complex observation[3].

The second major advantage is that the inherent low concentration of single-molecule techniques enables proteins, which are typically found *in vivo* at low concentrations, to be measured under their native conditions and even in living cells[4]. In these instances, ensemble-level techniques would typically require additional replication and purification. Similarly, some proteins aggregate at all but the lowest concentrations, here single-molecule methods enable interrogation of the monomeric species at equilibrium.

Finally, many single-molecule methods involve manipulation of individual molecules, this can facilitate probing the response to physical stimuli such as force or torque. Such measurements enable not only access to molecular mechanics[5], but also use of the molecules themselves as nanoscopic probes for sensing their nanoenvironment[6].

1.2 Single-molecule methods

In order to study a single molecule, it must be distinguished from its surroundings for detection. The first innovations in this came with the developments of fluorescence microscopy, which render the unlabelled background optically silent [7–9]. Other detection methods rely upon isolation of molecules from their surroundings. Several techniques rely upon tethering of a molecule to a specific site for sensing, such as to a micron-scale bead for optical/magnetic tweezers measurements [10,11], or to a cantilever tip for atomic force microscopy (AFM). Others spatially discriminate individual molecules using low molecular concentrations or small detection volumes, such as interferometric scattering microscopy (iSCAT) which has been used to detect low concentrations of molecules in nanochannels [12] or on surfaces [13], or nanopore technology which uses the change in ionic current as a single molecule translocates through a nanoscopic cavity for detection [14]. Finally, some methods rely upon the suppression of molecular motion through crystallisation or cryogenics, such as cryo-electron microscopy (cryoEM) whose sub-nanometer resolution can easily identify one molecule from another [15].

Once identified, molecules can be tracked to reveal co-localisation with other species, and hence indicate association and binding kinetics, such as in *in vivo* studies [16]. Diffusion measurements using either continuous imaging or spatial-temporal correlations can reveal insight into the size of molecules, as exemplified in fluorescence correlation spectroscopy (FCS) [17]. Additional perturbation of the molecule can allow inference of the molecular properties from measurement of the response to such perturbations. Methods including single-molecule förster energy resonance transfer microscopy (smFRET), single-molecule metal-induced energy transfer (smMiET) and fluorescence lifetime correlation spectroscopy (FLCS) use the photophysical response of fluorophores to incident photons to infer distances, orientations and more generally conformation and structure of molecules [18,19]. Meanwhile, mass photometry (MP) uses the scattering properties of a molecule to measure mass [20]. Nuclear magnetic resonance (NMR) uses the nuclear and electronic spin response to an applied magnetic field to determine chemical structure [21]. Single-molecule microfluidic diffusional sizing (smMDS) uses the diffusion across flow fields to determine molecular size [22,23]. Solid-state and protein nanopores use the impedance response of the molecule to an electric current to determine folding dynamics [24,25]. AFM uses the steric repulsion of a molecule to a cantilevered tip to directly read-out the conformation. Similarly, optical and magnetic tweezers can be used to probe the mechanical response of a molecule to applied forces and torques to measure the mechanical properties [5,26].

Some methods additionally rely upon the confinement of molecules for enhanced precision or measurement of additional time dependence. As a result various trapping methods have been developed which enable room temperature and solution phase trapping of molecules without surface tethering which has been reported to affect protein conformation[27,28]. The ABEL trap and later innovations use electric fields to confine molecules in an observation volume either directly[29,30] or via electro-osmotic flows[31], or to selectively filter individual molecules[32]. When combined with fluorescence microscopy, the former has been used for size measurements[33], and when integrated with smFRET has been used for conformation measurements[34,35]. Similarly, optical trapping has recently been applied to single molecules, either directly using the optical force induced from plasmonics[36], or through thermodiffusive forces induced from temperature gradients[37].

Geometric confinement has also been realised, which can facilitate similar trapping but without the need of externally applied fields, which has also been shown to affect protein folding[38]. Escape-time electrometry (ET*e*) uses the response of a molecule to differences in electric potential between areas of close and far separation from a charged surface to electrostatically confine molecules according to their effective charge[39]. Other geometric confinement methods involve reducing the dimensions of the volume available for molecules, either for enhanced sensitivity to diffusion measurements[40], or to effect longer-term confinement for single molecule imaging for up to minutes[41–44].

Of particular interest is escape-time electrometry, which focuses on not just the trapping of molecules in solution, but also of the measurement of molecular charge and size. Before the advent of ET*e*, molecular charge measurements had been somewhat neglected by the single-molecule community, despite the fact that many processes in nature rely upon electrostatic interactions. In addition to probing inter-molecular electrostatic interactions, charge measurements can also reveal a wealth of chemical and structural information, due in part to the complex nature of solution phase interactions between electrolytes and the molecule of interest, resulting in an "effective charge" of the molecule.

1.3 Effective electrical charge in solution

In the solution phase, molecules get their charge from the ionisation of their solvent exposed groups. We consider a weak acid HA, characterised by the dissociation, $\mathrm{HA} \rightleftharpoons \mathrm{H^+} + \mathrm{A^-}$, whose dimensionless acid dissociation constant K_a, is given by Eq. 1 where concentrations of species are denoted by square brackets relative to a standard concentration c°.

$$K_\mathrm{a} = \frac{\frac{[\mathrm{H^+}]}{c^\circ}\frac{[\mathrm{A^-}]}{c^\circ}}{\frac{[\mathrm{HA}]}{c^\circ}} \tag{1}$$

Such an acid, at a given pH, would have a degree of ionisation, α, as expressed in Eq. 2.

$$\alpha = \frac{[\mathrm{A^-}]}{[\mathrm{A^-}] + [\mathrm{HA}]} = \frac{1}{1 + 10^{z(\mathrm{pH}-\mathrm{p}K_\mathrm{a})}} \tag{2}$$

pK_a is the negative logarithm to base 10 of the acid dissociation constant, and z=-1 is the valency.

The structural charge of the entire molecule, q_str, is the most basic realisation of the ionisation of such groups, calculated by summing up the charge of all constituent groups at a given pH, as given in Eq. 3.

$$q_\mathrm{str} = \sum_i z_i \, \alpha_i \, e = \sum_i \frac{z_i e}{1 + 10^{z_i(\mathrm{pH}-\mathrm{p}K_{\mathrm{a},i})}} \tag{3}$$

p$K_{\mathrm{a},i}$ is the negative logarithm to base 10 of the acid dissociation constant for species i and $z_i = \pm 1$.

For most real molecules the effective charge $q_\mathrm{eff} \neq q_\mathrm{str}$ due to the contribution of two phenomena: charge regulation, and charge renormalisation[45].

1.3.1 Charge regulation

Charge regulation addresses the suppression of q_str of a particle due to local alterations in hydrogen ion concentration (and hence pH) in the vicinity of an electrostatic potential ϕ. Intuitively, in the case of a highly negatively charged surface we would expect an excess concentration of oppositely charged hydrogen ions present, thus locally decreasing the pH and hence reducing the degree of ionisation of a group with $\mathrm{p}K_\mathrm{a} \approx \mathrm{pH}$ (see Eq. 2). Assuming a Boltzmann distribution of hydrogen ions with reference to the bulk concentration $\left[\mathrm{H^+}\right]_\mathrm{b}$, we can express the surface concentration of

hydrogen ions, $\left[\mathrm{H}^{+}\right]_{\mathrm{s}}$, as in Eq. 4, with a dimensionless electrostatic potential at the surface ψ_{s}.

$$\left[\mathrm{H}^{+}\right]_{\mathrm{s}} = \left[\mathrm{H}^{+}\right]_{\mathrm{b}} \exp(-\psi_{\mathrm{s}}) \tag{4}$$

Ninham and Parsegian first showed that by equating $\left[\mathrm{H}^{+}\right]_{\mathrm{s}}$ in Eq. 4 to the local concentration of hydrogen ions determining the degree of ionisation of a group in Eq. 2, we recover the degree of regulated ionisation, α_{r}, given by Eq 5, with reference to the bulk pH, defined by $\left[\mathrm{H}^{+}\right]_{\mathrm{b}}$[46].

$$\alpha_{\mathrm{r}} = \frac{1}{1 + 10^{z_i(\mathrm{pH}-\mathrm{pK_a})} \exp\left(z_i \psi_{\mathrm{s}}\right)} \tag{5}$$

We note that for strongly acidic or basic groups, with $\mathrm{pK_a} \ll \mathrm{pH}$ or $\gg \mathrm{pH}$ respectively, or for weak surface potential, $\psi_{\mathrm{s}} \to 0$, we see no effect of charge regulation, $\alpha = \alpha_{\mathrm{r}}$, and therefore q_{str} would equal q_{eff} in the absence of charge renormalisation. We see this from combining Eq. 3 and 5.

1.3.2 Charge renormalisation

Charge renormalisation addresses the effective suppression of q_{str} of a particle due to counterions in solution, when observed from the far-field. Various descriptions of this phenomena exist in the literature, from the "Dressed-ion" to the "Counterion condensation" theory; they all revolve around the neutralising effects of high local concentrations of counterions in the vicinity of a charged particle[47,48].

An intuitive description of the phenomenon is that when the interaction energy between the particle and counterion is much larger than thermal energy, the counterion can be thought of as "condensed" onto the surface of the particle, neutralising some charge of the particle and leaving it unavailable as a freely diffusing ion. The result of this is that when observed from the far-field, the particle-counterion complex can now be considered as a new point object with apparent charge $q_{\mathrm{eff}} < q_{\mathrm{str}}$, which interacts with an electical potential ϕ with a free energy given by $F = q_{\mathrm{eff}}\phi$[49].

The magnitude of renormalisation, denoted $\zeta = \frac{q_{\mathrm{eff}}}{q_{\mathrm{str}}}$, has been shown to depend strongly on the 3D geometry of the charged object[50–52]. This we can think of in terms of the variable surface area to volume ratio of different shaped objects yielding variations in surface potentials, and hence different numbers of condensed counterions: a more compact object with low surface area to volume ratio we expect to have far higher surface potential, thus more counterion condensation and so a reduced ζ compared to a similar volume, less compact object.

1.3.3 Why measure effective electrical charge?

Because of the 3-dimensional shape dependency of charge renormalisation, measurements of q_{eff} can shed light on fine-grained properties of molecule structure[53,54], as well the conformational state of a molecule in solution, which can be otherwise hard to probe[55]. These are key handles for both analytics and diagnostics; the detection of a misfolded alpha-synuclein, as a biomarker of Parkinson's disease for example[56]. Effective charge can also be a useful analytical handle for the chemical state of a species, such as for the detection of small and charge-altering post-translational modifications (PTMs)[39,57].

More generally, effective charge is a key handle which governs the strength of electrostatic interactions, which underpin a host of biological processes[58] from binding kinetics and catalysis[59–61], to folding and stability[62–65], to aggregation and solubility[66,67]. Thus understanding a molecule's charge signature is key to understanding its role in such processes.

1.4 Escape-time electrometry (ET*e*)

This thesis focuses on the ET*e* technology. Chapter 2 outlines a series of applications of the technique as developed by Ruggeri and coworkers[39] - this method is described in the following sections. Chapters 3 and 4 describe developments of the technique and further methodology and theory is included in these chapters, where relevant.

1.4.1 Electrostatic trapping of molecules in solution

ET*e* employs a field-free fluid trap to confine molecules in solution[68]. The trap consists of physical indentations, so-called "pockets", in the top of nanoscale channels. The pockets and channels are etched into a microfluidic silicon dioxide (SiO_2) device sealed with a glass lid (SiO_2), a schematic of the device is shown in Fig. 1, and further details are outlined in Appendix A. The silanol groups on the channel surfaces, upon exposure to aqueous solution, hydrolyse to form negatively charged siloxide ions. Negatively charged molecules diffusing in the channels are repelled by the same-charged surface groups and so find an electrostatic energy minima inside a pocket, the electrostatic trap.

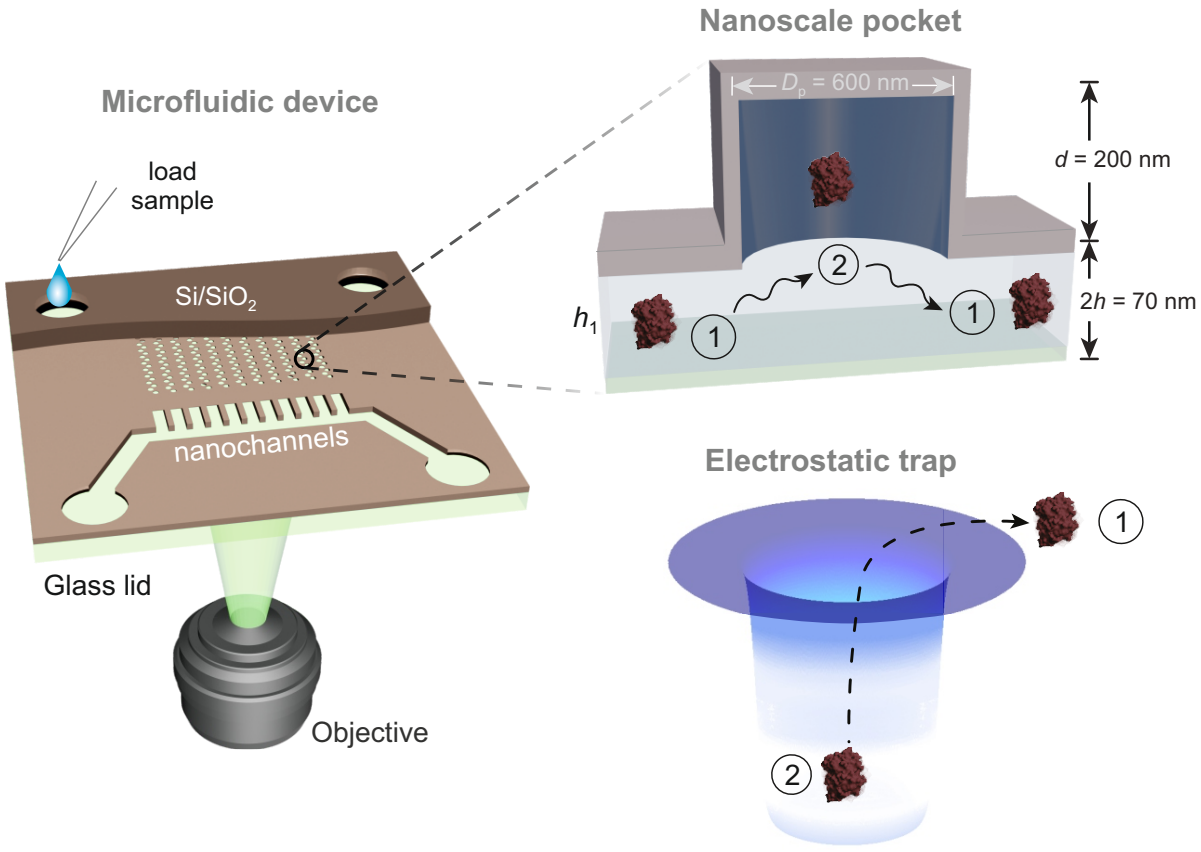


Fig. 1 **Schematic of the field-free fluid trap in an ET*e* device** A microfluidic device is fabricated with 10s of parallel nanochannels, each with 1000s of pockets in. Each pocket creates an individual electrostatic trap to confine molecules.

To evaluate the strength of the electrostatic trap we consider the electrostatic potential, ϕ, whose distribution in the vicinity of a charged surface is given by the Poisson-Boltzmann (PB) equation, Eq. 6[69]. For simplicity we introduce a dimensionless electrostatic potential, ψ, related to ϕ by the elementary charge, e, and thermal energy, $k_B T$, according to $\psi = e\phi/k_B T$.

$$\nabla^2\psi = \kappa^2 \sinh\psi. \tag{6}$$

κ is the inverse Debye length, which describes the screening length of electrostatic interactions in an aqueous medium with dielectric constant ε and monovalent salt concentration c, as given by Eq. 7.

$$\kappa^{-1} = \sqrt{\frac{\varepsilon\varepsilon_0 k_B T}{2cN_A e^2}}. \tag{7}$$

ε_0 is the dielectric constant of a vacuum and N_A is Avogadro's number. Applied to the fluid trap geometry of approximately two parallel plates with electrostatic potential at their surfaces ϕ_s separated by a distance z_{max}, Eq. 6 becomes the linearised PB equation, $\nabla^2\psi = \kappa^2\psi$, which yields a solution, valid far away from the surfaces, or for small ϕ_s, as given in Eq. 8, and displayed in Fig. 2.

$$\phi(z) = 2\,\phi_s \cosh\left[\kappa\left(z - \frac{z_{max}}{2}\right)\right] \tag{8}$$

The energy well depth, W, of the trap for a point particle of effective charge q_{eff} is evaluated by considering the difference in free energy of the system when the particle is located in the channel (state 1), compared to in the pocket (state 2). One major constituent of W is due to the local electrostatic potential minima inside the pocket. We define the electrostatic interaction free energy, ΔF_{el}, as the difference in electrostatic free energies $F_{el}(z) = q_{eff}\phi(z)$, when the particle is positioned at the midplane of the channel, with $z{=}h$, and at the midplane of the pocket, with $z{=}h{+}d/2$, as given in Eq. 9 and displayed in Fig. 2[39].

$$\Delta F_{el} = F_{el}|_1 - F_{el}|_2 = q_{eff}(\phi(h) - \phi(h + d/2)). \tag{9}$$

By design we ensure that d and κ are sufficient such that $\phi(h + d/2) \approx 0$, therefore by combining Eq. 8 and 9 we obtain an expression for the electrostatic contribution to the well depth according to Eq. 10, where we have defined the potential at the midplane of the channel as the midplane

potential, $\phi_{\mathrm{m}} = \phi(h)$.

$$W \propto \Delta F_{\mathrm{el}} = q_{\mathrm{eff}}\, \phi_{\mathrm{m}} = 2\, q_{\mathrm{eff}}\, \phi_{\mathrm{s}} \exp(-\kappa h) \tag{10}$$

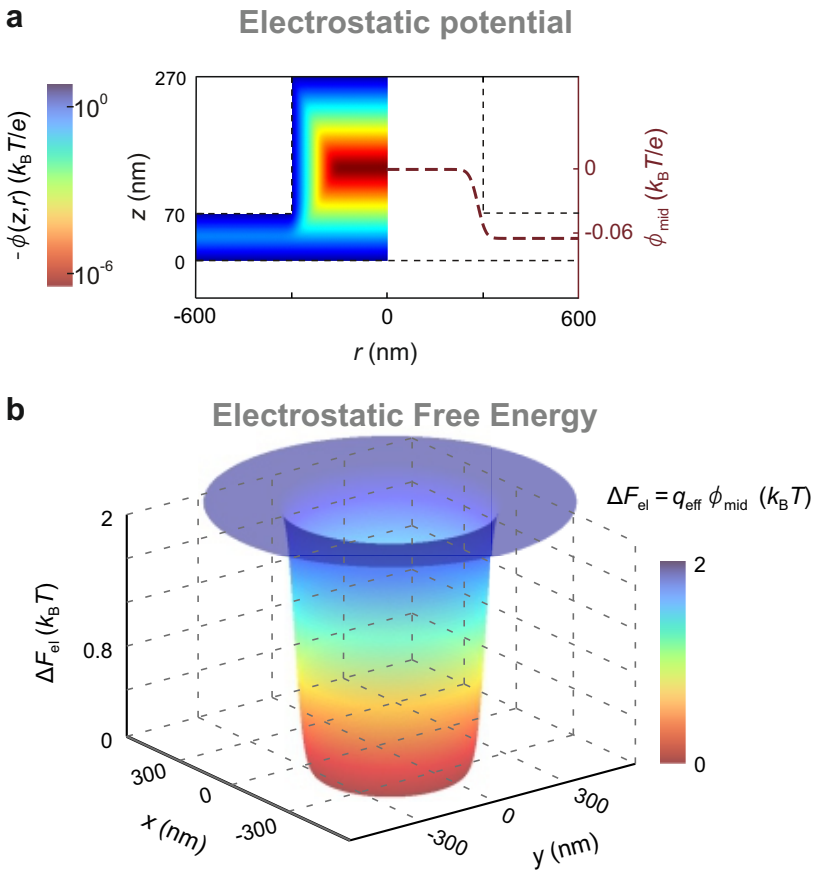


Fig. 2 **Visualisation of the electrostatic potential and free energy landscape of an ETe trap** (**a**) The electrostatic potential, $\phi(z)$, is calculated by solving Eq. 6 using COMSOL mulitphysics. The minimum potential at each radial coordinate is extracted, ϕ_{mid}. ϕ_{mid} in the pocket is zero by design, and is defined as ϕ_{m} in the channel. (**b**) The electrostatic free energy difference along the midplane of the device relative to the pocket region is calculated as $\Delta F_{\mathrm{el}} = q_{\mathrm{eff}}\, \phi_{\mathrm{mid}}$.

1.4.2 Escape from the electrostatic fluid trap

From Eq. 10 we can see that the energy well depth of the trap can be tuned by altering one of ϕ_{s}, κ or h, so that W is similar to thermal energy[68]. This means that a trapped particle will escape stochastically from the potential well on a reasonable timescale (milliseconds to seconds). It is a measurement of this time for a particle to escape, Δt, which ETe is based upon.

To evaluate the expected trapping time of a particle we consider the relation derived by Kramers for the rate of migration of particles over a **high** potential barrier[70]. Rearranged for our system we can express the typical escape time, t_{esc}, for a point particle out of an energy well

of depth W, according to Eq. 11.

$$t_{\mathrm{esc}} = t_{\mathrm{r}} \exp\left(\frac{W}{k_{\mathrm{B}}T}\right) \tag{11}$$

t_{r} is the position relaxation time of the particle, which we expect to be well approximated by the average time taken for a particle to diffuse to the edge of the trap (since $t_{\mathrm{r}} = t_{\mathrm{esc}}$ for $W = 0$). For a particle with translational diffusion coefficient D_{t}, trapped in a pocket of diameter D_{p}, we expect $t_{\mathrm{r}} \approx \frac{{D_{\mathrm{p}}}^2}{4D_{\mathrm{t}}}$.

1.4.3 Measurement of q_{eff} using ETe

Measurement of t_{esc}

Fluorescently labelled molecules are loaded into the pocket array in ETe devices through Poiseille flow, generated using a vacuum pump to impose a pressure difference along the length of the channels. Once loaded, flow is arrested and the motion of molecules diffusing in the trapping structure is observed using a widefield fluorescent microscope, details of which are given in Appendix B. Escape events of molecules leaving from individual traps after a residence time Δt are measured from analysis of the intensity-time traces from signals collected from each pocket in the field of view (FOV), see Fig. 3. Further details of the movie analysis is given in Appendix C. The escape process is a Poisson process meaning that Δt is exponentially distributed. A histogram of Δt is constructed and fit with an exponential distribution of the form $P(\Delta t) = E \exp\left(\frac{-\Delta t}{t_{\mathrm{esc}}}\right) + \mathrm{c}$. Where escapes occur at multiple timescales, t_i, multi-exponential fitting of the histogram of the form given by Eq. 12 is done.

$$P(\Delta t) = \sum_i E_i \exp\left(\frac{-\Delta t}{t_i}\right) + \mathrm{c} \tag{12}$$

Calculation of W from t_{esc}

By design $W \gtrsim k_{\mathrm{B}}T$ in practice, and so Eq. 11 does not hold exactly. In order to precisely examine how t_{esc} depends on W we employ Brownian Dynamics (BD) simulations of particles diffusing in the trap landscape, see Appendix D. By comparing simulated t_{esc} for a particle of diffusion coefficient D_{t}, in a trap of well depth W, with experimentally measured t_{esc} for a molecule with the same diffusion coefficient, we can map the measured t_{esc} to a W experienced

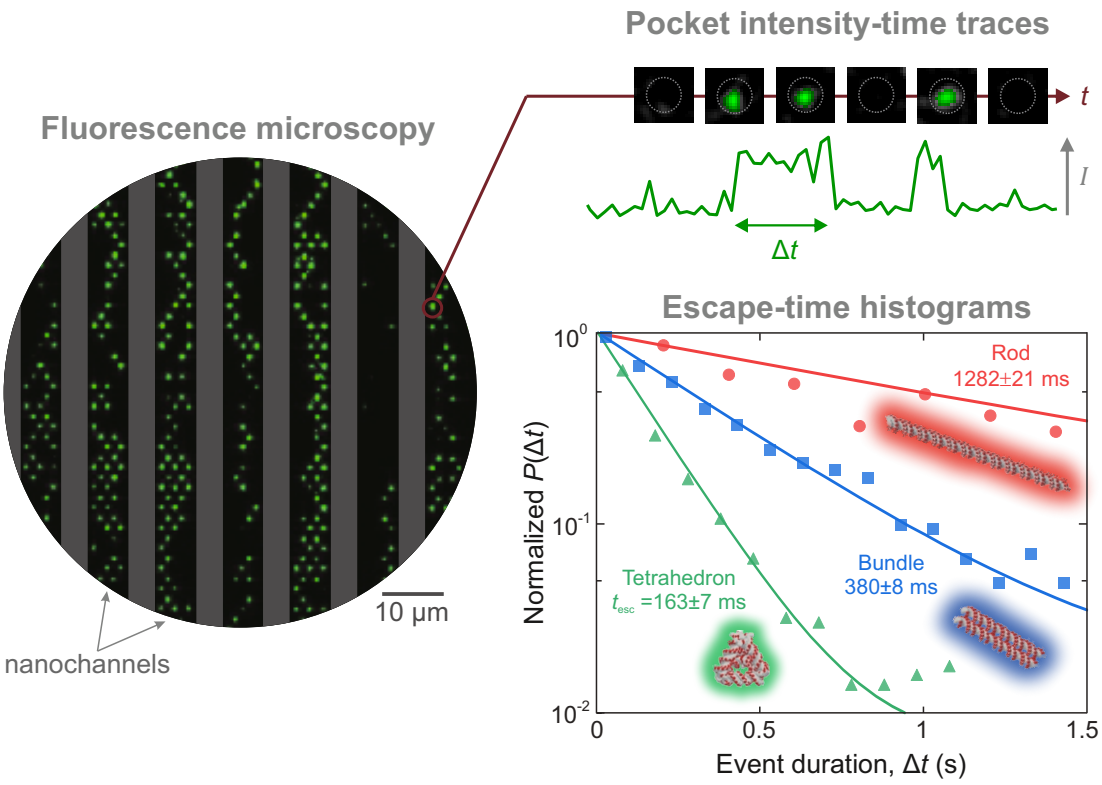


Fig. 3 **Schematic of an ET*e* experiment** Widefield fluorescence microscopy is used to simultaneously image 1000s of pockets. Residence times Δt from individual escape events are recorded from the intensity-time traces from each pocket. Histograms of residence times are constructed and fit with Eq. 12.

by the molecule. We find in general that for $W > 4k_{\mathrm{B}}T$, t_{esc} follows the exponential relation to W as given by Eq. 11[71].

Dual-focus fluorescence correlation spectroscopy (2f-FCS) is used to independently measure the D_{t} of the molecule. This was chosen over other diffusion measurements such as dynamic light scattering (DLS)[72] or traditional fluorescence correlation spectroscopy due to the high precision and absolute nature of the technique without the need for reference measurements[73]. The measurement procedure is outlined in the supporting information of the paper in section 2.1.1[74], see Appendix G.

Calculation of q_{eff} from W

The effective charge, q_{eff}, is the global quantity of interest in ET*e* measurements, and is defined by the constant of proportionality between the electrostatic interaction free energy, ΔF_{el}, and the midplane potential, ϕ_{m}, as in Eq. 10. ΔF_{el} fully describes the well depth for point particles confined to the midplane of the channel and pocket, however thermal energy ensures that fluctuations in the z-direction are always present, and thus we include a fluctuation free energy f, to the full expression for W, as shown in Eq. 13. This fluctuation free energy can be assigned

to both the translational departure of the molecule from the midplane, ΔF_{trans}, and also the spatial departure of parts of a finite-size molecule from the midplane while rotating, ΔF_{rot} [75]. The relative contributions of the fluctuation terms are given in Fig. 4.

$$W = \Delta F_{\text{el}} + f = \Delta F_{\text{el}} + \Delta F_{\text{trans}} + \Delta F_{\text{rot}} \tag{13}$$

Calculation of ΔF_{trans} and ΔF_{rot} are therefore required to extract q_{eff} in a given measurement. ΔF_{trans} is calculated from considering the full electrostatic free energy difference (Eq. 14) between a point particle in the channel compared to in the pocket, calculated from integration of the partition function for a point particle freely diffusing in the z-direction, Z (Eq. 15).

$$q_{\text{eff}}\,\phi_{\text{m}} + \Delta \text{F}_{\text{trans}} = -k_{\text{B}}\,T \ln Z|_1 + k_{\text{B}}\,T \ln Z|_2 \tag{14}$$

$$Z = \int_0^{z_{\text{max}}} \exp\left(\frac{-F_{\text{el}}(z)}{k_{\text{B}}\,T}\right) \text{d}z \tag{15}$$

Where $F_{\text{el}}(z) = q_{\text{eff}}\phi(z)$ is the electrostatic free energy when the particle is located at a height z. By solving Eq. 14 using Eq. 10 and Eq. 15 for the limits of $z_{\text{max}} = 2h$ and $z_{\text{max}} = (2h + d)$ for the channel and pocket state respectively, we obtain ΔF_{trans}. We note that in the limit of no electrostatics where $\phi_{\text{m}} = 0$ and $\Delta F_{\text{el}} = 0$, $\Delta F_{\text{trans}} = k_{\text{B}}\,T \ln\left(\frac{2h+d}{2h}\right)$; there is still a non zero energy well depth of the trap, the "entropic well depth". This is the case explored by experiments in Chapter 4.

ΔF_{rot} contributes only a small fraction of the total well depth, $\lesssim 5\%$ for 60 base-pair (bp) DNA generally, so is only included in the calculation for molecules larger than 60 bp DNA, such as the measurements in section 2.1. Details of the calculation is outlined in the supporting information of the paper in section 2.1.1 [74], see Appendix G.

It is apparent from Eq. 15 and 10 that in order to calculate f a series of physical parameters are required. *a priori* we can determine both κ and h: κ through conductance measurements of the analyte converted to salt concentration, and h from AFM scans of the device. We cannot *a priori* know ϕ_{s} and q_{eff} - indeed the latter would be rather cyclical, however we note that because of the weak, linear dependence of f on q_{eff} (weak compared to the exponential dependence of ΔF_{el}), an initial estimate of q_{eff} suffices. ϕ_{s} is obtained using a calibration measurement of a molecule whose theoretical q_{eff} has been explicitly calculated; a ϕ_{s} value is selected so that the measured and calculated q_{eff} values agree. Details of this calculation is outlined in the supporting

information of the paper in section 2.1.1[74], see Appendix G. Thus using an initial estimate for q_{eff}, f is calculated, and hence a better estimate for q_{eff} is obtained, and this is repeated iteratively until a final measured value is obtained, q_{m}.

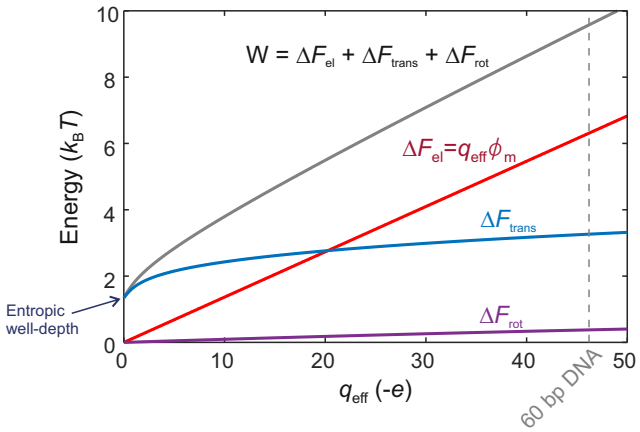


Fig. 4 **Calculation of contributions to the well depth**. ΔF_{el}, $\Delta F_{\mathrm{trans}}$, and $\Delta \mathrm{F}_{\mathrm{rot}}$ are calculated for a line charge of length 20.4 nm (corresponding to 60 bp DNA) with varying charge in 1 mM NaCl, in a geometry corresponding to $2h$=70 nm, d=200 nm, with $\phi_{\mathrm{s}} = -2.7\, k_{\mathrm{B}}T/e$. We note that for an uncharged object there is still a considerable well depth - the entropic well depth.

1.5 Thesis outline

This thesis identifies several applications and recent developments of the escape-time electrometry technology (ET*e*), and consists of 3 sections:

- **Chapter 2** illustrates ensemble ET*e* measurements of 3D molecular conformation and the detection of post-translational modifications. This chapter contains a manuscript published in a peer-reviewed journal[74].

- **Chapter 3** exemplifies the single-molecule nature of ET*e* with the introduction of trajectory-based ET*e* (trET*e*) and a new method called flow-based ET*e* (FlET*e*).

- **Chapter 4** describes ET*e* measurements under physiological conditions and the advent of escape time size-and-shape spectrometry (ETsy). This chapter contains a manuscript submitted to a peer-reviewed journal.

1.6 List of publications

List of publications included in this thesis

- Far-Field Electrostatic Signatures of Macromolecular 3D Conformation. G. Kloes†, **T. J. D. Bennett**†, A. Chapet-Batlle, A. Behjatian, A. J. Turberfield, and M. Krishnan. *Nano Letters*, 22(19):7834-7840, 2022.

- Single-molecule size-and-shape spectrometry on a chip. X. Zhu†, **T. J. D. Bennett**†, K. C. Zouboulis, D. Soulias, M. Gryzbek, J. L. P. Benesch, A. H. El-Sagheer, Ü. Coskun, and M. Krishnan. *Submitted*, 2024

† Contributed equally

List of publications not included in this thesis

- Sensing the structural and conformational properties of single-stranded nucleic acids using electrometry and molecular simulations. R. Walker-Gibbons†, X. Zhu†, A. Behjatian, **T. J. D. Bennett**, and M. Krishnan. *Sci Rep*, 14, 20582, 2024.

† Contributed equally

Chapter 2

Ensemble ET*e*

The motivation for measurements of effective charge in solution is clear from section 1.3.3. This chapter outlines two such use cases with proof of principle measurements: measurement of 3D conformation in section 2.1, and detection of post-translational modifications in section 2.2.

2.1 Measurements of 3D molecular conformation

As mentioned in section 1.3, the effective charge of a molecule is sensitive to the 3D conformation through renormalisation. In order to probe the sensitivity of ET*e* to conformation, a series of DNA nanostructures, chosen due to their low pK_a charged groups so as to minimise charge regulation, were measured. These experiments are described by the following paper: "Far-field electrostatic signatures of macromolecular 3D conformation"[74].

2.1.1 Far-Field Electrostatic Signatures of Macromolecular 3D Conformation

Briefly, we show that DNA nanostructures, which are characterised by approximately identical structural charge and diffusion coefficients, yield highly variable effective charge values. These measurements illustrate the dependence of q_{eff} on conformation, and the ability of ET*e* to distinguish such differences. In addition, from comparison with calculations of theoretical q_{eff} values, we show that fine-structural parameters can be inferred from such measurements, such as the interhelical spacing in the square-tile nanostructure.

Such proof of concept measurements of 3D conformation stand to facilitate further studies into conformation states of biomolecules in solution using ET*e*. Combined with the advancements made in single-molecule ET*e* as outlined in Chapter 3, we envisage investigation of the conformation states of RNA molecules in heterogeneous samples, which are notoriously difficult to probe[55].

The supporting information for this paper can be found in Appendix G.

NANO LETTERS

pubs.acs.org/NanoLett

Letter

Far-Field Electrostatic Signatures of Macromolecular 3D Conformation

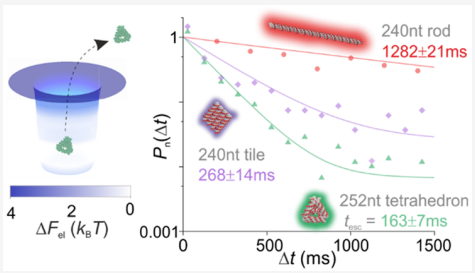
Gunnar Kloes,[§] Timothy J. D. Bennett,[§] Alma Chapet-Batlle, Ali Behjatian, Andrew J. Turberfield, and Madhavi Krishnan*

Cite This: *Nano Lett.* 2022, 22, 7834–7840

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: In solution as in vacuum, the electrostatic field distribution in the vicinity of a charged object carries information on its three-dimensional geometry. We report on an experimental study exploring the effect of molecular shape on long-range electrostatic interactions in solution. Working with DNA nanostructures carrying approximately equal amounts of total charge but each in a different three-dimensional conformation, we demonstrate that the geometry of the distribution of charge in a molecule has substantial impact on its electrical interactions. For instance, a tetrahedral structure, which is the most compact distribution of charge we tested, can create a far-field effect that is effectively identical to that of a rod-shaped molecule carrying half the amount of total structural charge. Our experiments demonstrate that escape-time electrometry (ET*e*) furnishes a rapid and facile method to screen and identify 3D conformations of charged biomolecules or molecular complexes in solution.



KEYWORDS: *Electrostatics, electrometry, DNA nanostructures, molecular interactions, 3D conformation*

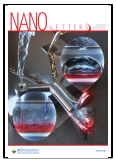
Downloaded via UNIV OF OXFORD on September 26, 2024 at 13:08:25 (UTC).
See https://pubs.acs.org/sharingguidelines for options on how to legitimately share published articles.

The structures of materials, phase transitions, chemical reactions, and biological processes fundamentally rely on molecular interactions. Such interactions occur across a range of length scales from short-range bond formation and breakage to long-range Coulombic interactions that stem from the presence of a net molecular electrical charge. The decisive role of three-dimensional molecular shape in short-range interactions and reactive processes is familiar from a wealth of examples provided by stereochemistry and chemical and biochemical selectivity. But molecular shape can also have a strong and measurable impact on the long-range electrostatic interaction.

Consider an isolated electrically charged object in the fluid phase. It is well-known that geometry affects the distribution of electrical potential around the object, in turn influencing the local distribution of counterions.[1−3] Ion density distributions around a charged object in solution depend significantly on whether the object is a sphere, rod, or a plane, for instance.[4−8] The influence of object geometry on counterion distribution and dynamics may manifest in a variety of ways in experiment. Some examples of experimental observables include the mobility of a charged object in an electrical field, the osmotic pressure of a suspension of charged particles, ionic strength dependent polyelectrolyte properties such as persistence length as well as in intra- and intermolecular interaction energies and forces.[5,8−12] Detection of the electrical field or potential distribution at a fixed distance from an isolated test object, using interaction energy measurements, could therefore reveal information not only on the total amount of charge the object carries but also on the details of the 3D distribution of charge.[10] Importantly, for uniformly charged entities carrying identical amounts of total charge, any differences in interaction energy, measured under identical conditions, should stem solely from the shape of the object.

A significant body of theoretical work discusses the interplay between the geometry of a charged object and its electrostatic properties such as the potential distribution in the surrounding electrolyte and the potential of mean force or interaction energy with another charged entity.[1−7] These studies generally focus on a parameter called the charge renormalization factor, $\eta = \frac{q_{eff}}{q_{str}}$, that relates the structural charge of the object, q_{str}, with its effective charge, q_{eff}. In the molecular context, q_{str} denotes the net electrical charge in the molecular structure and is given by the sum of charges carried by ionized structural groups and any bound ions from the electrolyte. The effective charge, q_{eff}, may depart from the nominal structural value for a variety of reasons. In general charge renormalization theories predict $\eta \rightarrow 1$ for weakly charged objects where the average spacing between charges is much larger than the Bjerrum length, l_B, for charged cylinders or $\sqrt{Rl_B}$ for charged spheres of

Received: June 23, 2022
Revised: September 7, 2022
Published: September 20, 2022



ACS Publications

© 2022 The Authors. Published by American Chemical Society

radius R.[1,5,6] The Bjerrum length $l_B = \frac{e^2}{4\pi\varepsilon\varepsilon_0 k_B T} \approx 0.7$ nm in water is the separation at which the energy of interaction between two elementary charges (e) is equal to k_BT, where k_B is the Boltzmann constant and T is the absolute temperature. On the other hand, for objects with high charge densities we expect $\eta < 1$.[1] The role of charge renormalization is further accentuated under conditions of weak electrostatic screening, with η generally decreasing as the salt concentration decreases.[5,6,8] Importantly, since η depends on the 3D distribution of charge in an object, q_{eff} depends strongly on the shape of the object.[4]

Although the renormalization factor has most often been discussed in the context of the effective charge of an isolated charged object, we have recently found that the value calculated for an isolated object turns out to be quantitatively very similar to the effective charge that determines the electrical free energy in a molecular interaction (between the object of interest and another charged entity).[5,6,8,13,14] Our previous experimental work on a range of different biomolecules has further demonstrated that measurements of effective charge, as inferred from the interaction of a charged molecule with a like-charged flat surface in solution for example, correspond remarkably well to q_{eff} values calculated using Poisson−Boltzmann (PB) interaction free energies.[8,10] Further, it is worth noting that our interaction based definition of effective charge is the same as that used in Kjellander's "dressed ion" theory.[13,14]

In this study we use DNA nanostructures to examine the impact of 3D macromolecular conformation on long-range molecular electrostatic interactions and, conversely, to explore the possibility of using high-precision electrostatics measurements to shed light on 3D macromolecular conformation and structure.[15] We measure electrostatic interaction energies and corresponding effective charge values for various DNA constructs confined between two parallel charged plates at a separation $2h$. We work in the regime given by $\kappa h \approx 4-8$, i.e., where the typical separation, h, between molecule and probe is greater than the Debye length, κ^{-1}. Here $\kappa^{-1} = (8\pi l_B N_A I)^{-1/2}$ is a length scale which characterizes the rate of exponential decay of the electrical potential with distance from the surface of a charged object, where N_A is Avogadro's number and I is the ionic strength of the electrolyte ($\kappa^{-1} = 4-9$ nm in this study).

Here, we consider DNA nanostructures designed to carry effectively the same total structural charge, q_{str}, but of four different molecular shapes representing conformations of varying "compactness", as reflected in their respective radii of gyration, R_g, the root-mean-square distance of all the atoms in the molecule from its center of geometry (Table S5). Beginning with a 120 bp approximately cylindrical double-helix ($q_{str} = -238$ e; $R_g = 12$ nm), we consider progressively more compact molecular 3D conformations given by a 2-helix bundle ($q_{str} = -237$ e; $R_g = 6$ nm), a square-tile ($q_{str} = -237$ e; $R_g = 4$ nm), and finally a DNA tetrahedron ($q_{str} = -250$ e; $R_g = 3.75$ nm)[16] (Supporting Information section 1). For reference, we also examine a 60 bp DNA cylindrical double-helix bearing half the amount of structural charge carried by the other constructs ($q_{str} = -120$ e; $R_g = 6$ nm).

MEASURING THE EFFECTIVE CHARGE OF DNA NANOSTRUCTURES USING ESCAPE-TIME ELECTROMETRY (ET*e*)

We performed electrostatic interaction energy measurements on the DNA nanostructure species, labeled with one ATTO 532 fluorophore per nanostructure, using our recently developed escape time electrometry (ET*e*) technique[10,17−19] (Figure 1). ET*e* offers high precision measurements of electrostatic interaction energies and therefore of molecular effective charge in the fluid phase. Briefly, we used wide-field fluorescence microscopy to image the escape dynamics of single molecules, loaded into a system of multiple parallel slits each containing lattices of traps as shown in Figure 1. We first extracted precise measurements of the average molecular time to escape, t_{esc}, (Figure 1b,d) and then converted t_{esc} to electrostatic interaction free energies, ΔF_{el}. Next, we deduced measured effective charge values, q_m, as in previous work[10,17−19] (Figure 1a). A full description of the details of the ET*e* measurement is provided in the Supporting Information section 2.

In our work, the relationship between ΔF_{el} and molecular effective charge, q_{eff}, is given by $\Delta F_{el} = q_{eff}\Delta\phi_{mid}$, where ϕ_{mid} is the electrical potential at the midplane of a parallel-plate system, and Δ denotes the difference between two states of the molecule where it is outside and inside the trap region (also referred to as the "slit" and "pocket" states), as shown in Figure S8a of the Supporting Information. This can be simplified to $\Delta F_{el} = q_{eff}\phi_m$, where $\phi_m \approx \Delta\phi_{mid} \approx 2\phi_s \exp(-\kappa h)$ is the electrical potential in the nanoslit region of height $2h$, since ϕ_{mid} in the pocket region is zero by design (Figure 1c). Here, ϕ_s is the effective surface electrical potential at the silica walls of the slit, and κh is a dimensionless system size or a dimensionless separation distance between a molecule and a slit surface. It follows that

$$\ln\frac{\Delta F_{el}}{k_BT} = \ln\frac{2q_{eff}\phi_s}{k_BT} - \kappa h \quad (1)$$

The two unknowns in eq 1 are q_{eff}, which describes the molecule of interest, and ϕ_s which describes the measurement device. In order to determine the value of ϕ_s in a given experiment, we first performed a calibration measurement using a well-characterized molecular species of known effective charge under experimental consitions, namely 60bp double-stranded DNA fragment carrying two fluorophore labels characterized by $q_{eff} = -46.2$ e.[10,17,19] Values for ϕ_s obtained in four different measurement devices are noted in Figure 1e and are consistent with previously measured values.[10,17−19] Measurements of t_{esc} under various conditions of monovalent salt concentration ($c = 1.1-5.5$ mM; $\kappa^{-1} = 4-9$ nm) and slit height are converted to electrostatic free energies, ΔF_{el}, which are plotted against κh. A fit to eq 1 in conjunction with the value for ϕ_s yields measured values of q_{eff} for each molecular species, as discussed in detail in Supporting Information section 2.5. Measured effective charge values, denoted as q_m are displayed in Figure 1e, displaying constant behavior over the range of κh probed in experiment. Although q_{eff} is expected to depend on salt concentration, we do not expect significant variation over the narrow range of c probed in these experiments (Supporting Information section 3.3).[5,6,8] Finally, we work with a measurement uncertainty of ≈2−7% on t_{esc} and on the molecular effective charge (Supporting Information

https://doi.org/10.1021/acs.nanolett.2c02485
Nano Lett. 2022, 22, 7834−7840

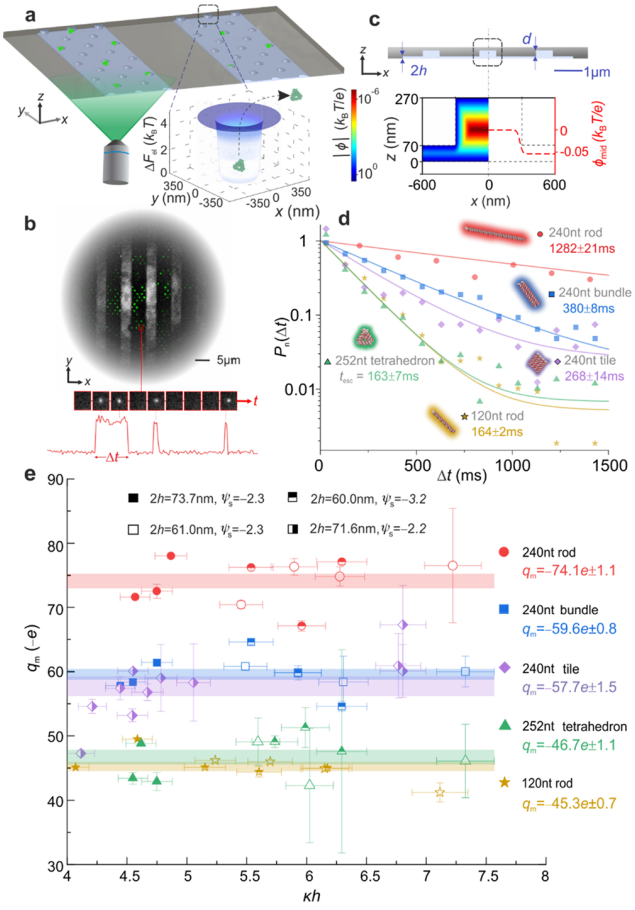


Figure 1. Measurement of effective charge of DNA nanostructures using ET*e*. (a) Graphic representation of an ET*e* device consisting of multiple parallel fluidic slits housing arrays of electrostatic fluidic traps, imaged using wide-field fluorescence microscopy. Calculated representative electrostatic energy landscape (ΔF_{el}) for a single trap (inset). (b) Maximum intensity projection of the fluorescence signal recorded over a time period of 20s in a typical ET*e* experiment (top). Temporal evolution of the fluorescence signal in a representative single trap displaying a series of images and the corresponding intensity trace (bottom). (c) Graphical representation of the cross-section of a fluidic slit of height $2h \approx 70$ nm and pocket depth $d \approx 200$ nm (top). Below: Calculated electrostatic potential distribution, ϕ, presented on logarithmic scale for a salt concentration $c = 1.5$ mM and $h = 35$ nm (left half-space) and line plot of midplane potential value, ϕ_{mid}, in the vicinity of a single pocket (right half-space), displaying a value $\phi_m \approx -0.05 k_B T/e$ in the slit region, corresponding to an effective surface potential of $\phi_s = -2.3 k_B T/e$. (d) Probability density distributions of residence times, obtained from 10 to 20 min of imaging, are fitted to the expression $P(\Delta t) \propto \exp(-\Delta t/t_{esc})$ to obtain an average escape time, t_{esc}, for each nanostructure species. $P(\Delta t)$ distributions are presented such that the maximum value of the fitted curve is set to 1 in order to foster visual comparison of t_{esc} for all species measured. t_{esc} for species measured under identical conditions may be compared directly. A factor 8 difference in t_{esc} values between the tetrahedron ($t_{esc} = 163 \pm 7$ ms) and the 120 bp rod ($t_{esc} = 1282 \pm 21$ ms) implies a substantially smaller ΔF_{el} value for the former. Note that in general we do not include in the fitting process 1−2 data points at the shortest measured lag times, Δt. This is because escape events recorded at the shortest durations can contain contributions from transiently trapped, weakly charged molecular species such as free dye, degraded material, or free DNA strands in the background solution. All measurements were performed in devices with slits of depth $2h = 71.6$ nm and surface potential value $\phi_s = -2.3 k_B T/e$, under similar electrolyte conditions ($\kappa h \approx 4.5$). Measurements on the square-tile, however, were performed in a similar device but with $2h = 73.7$ nm and $\phi_s = -2.2 k_B T/e$, yielding a t_{esc} value that is not directly quantitatively comparable with values for the other species shown. (e) q_m values inferred for each individual measurement for all species are constant within measurement uncertainty for a given nanostructure species, over a range of salt concentrations ($c = 1.1$−5.5 mM) as well as across different measurement devices where slit heights ranged from $2h \approx 60.3$ to 73.7 nm and dimensionless surface potentials ranged from $\psi_s = \frac{e\phi_s}{k_B T} = -3.2$ to -2.2. Also noted are q_m values averaged across the range of κh probed.

section 2.6), which is sufficient to delineate broad trends arising from molecular geometry.

CALCULATED EFFECTIVE CHARGE VALUES FOR DNA NANOSTRUCTURES

Despite the fact that all our test DNA nanostructures carry nearly identical amounts of the structural charge, we note a gradual reduction in the magnitude of q_m of nearly a factor of 2 from the most extended nanostructure (with the largest value of R_g), the 120 basepair double-helix, to the most compact, the tetrahedron. We further calculated theoretically expected values of the effective charge of the DNA nanostructures using an approach described previously[8] and outlined in detail in the Supporting Information section 3. Individual nanostructures were modeled as an assembly of smooth charged cylinders each representing a segment of double helical DNA. The cylinders have a radius $r_{cyl} = 1.2$ nm, a length corresponding to a rise per base pair value of 0.34 nm, and carry a uniform charge density $\sigma = q_{str}/A$, where A is the total surface area of the structure.[31] Distributions of dimensionless electrical potential $\psi = \frac{e\phi}{k_B T}$ on the surfaces of the nanostructures (ψ_s) as well as in the surrounding electrolyte in the vicinity of the nanostructure (ψ) are displayed in Figure 2. We note remarkable agreement in general between the experimental q_m values and the calculated q_{eff} values in each case (Figures 2 and 3). We further found that a tetrahedral arrangement of a total electrical charge of $q_{str} = -250\ e$ ($R_g \cong 4$ nm) is in fact nearly indistinguishable in the present measurement from a linear arrangement of half that amount of charge, as encountered, for example, in a fragment of 60 bp DNA ($q_{str} = -122\ e$; $R_g \cong 6$ nm). We did however note a 10% discrepancy between the measured and calculated effective charge values for the 120 bp double-helix (Figure 2). This disparity may arise from the fact that whereas we model the double-helix as a rigid cylindrical rod, the 120 bp fragment of DNA in our experiments may be less extended as it is assembled from four oligonucleotides (see Supporting Information section 1), possibly resulting in a structure that is more flexible than a canonical two-strand double-helix.

In order to illustrate the ability of a measurement of q_{eff} to provide information on molecular 3D geometry, we mapped the nanostructure electrostatics problem onto that of charge-carrying spheroids of fixed total volume V and total charge $q_{str} = -240\ e$. We set the volume of each spheroid $V = \frac{4\pi a^2 c}{3}$ to be equal to that of a spherical shell of radius equal to the radius of gyration of the DNA tetrahedron ($R_g \approx 4$ nm), our most conformationally compact nanostructure. We then varied the semiaxis lengths a and c in order to capture the range of

https://doi.org/10.1021/acs.nanolett.2c02485
Nano Lett. 2022, 22, 7834−7840

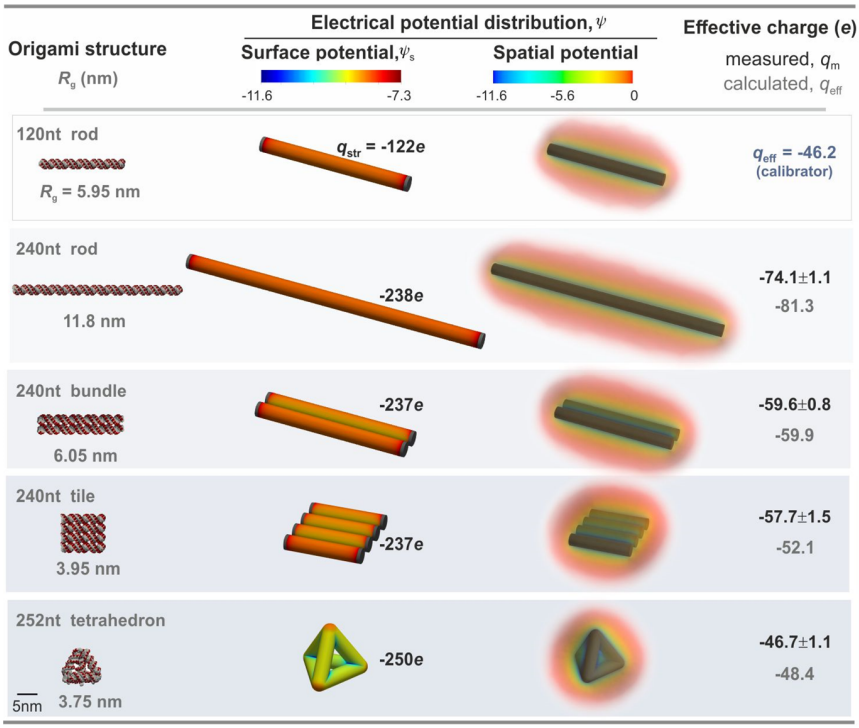


Figure 2. Electrostatic modeling and effective charge measurements for DNA nanostructures. Columns 1 and 2 list values for the radius of gyration, R_g (computed for all atoms in the molecular structures), DNA nanostructure size in nucleotides (nt), and total structural charge, q_{str}, including the contribution of the fluorescent dye. Electrostatic modeling of the nanostructure was performed by solving the PB equation for coarse-grained cylinder-based models of the nanostructures immersed in an electrolyte containing 1.5 mM NaCl (described further in Supporting Information section 3). Distributions of dimensionless surface and spatial electrical potential, ψ_s (column 2) and ψ (column 3), respectively, are calculated for each nanostructure model. Spatial distributions of electrical potential, ψ, correspond to the local electrolyte volume surrounding the nanostructures. ψ_s and ψ are presented on an identical scale of electrical potential for all nanostructures and emphasize the sensitivity of the spatial potential distribution around the object to 3D molecular conformation and "compactness". Column 4 compares calculated effective charge values, q_{eff} (bottom), with measured effective charge values, q_m (top), that are averages over the series of ET*e* measurements presented in Figure 1e.

nanostructure shapes in the study. We then calculated effective charge values for hollow, uniformly charged spheroids representing each nanostructure. As reflected in the PB equation (Supporting Information S3), charge renormalization is a manifestation of nonlinearity in the underlying electrostatics. A low-density charge distribution generally places the electrostatics problem in the linear regime, characterized by $|\psi| < 1$, where $q_{eff} \rightarrow q_{str}$. A more compact or dense charge distribution would generally imply a large value of $|\psi|$, particularly in the vicinity of the object, and we expect $q_{eff} \ll q_{str}$ as a consequence[5,8] (Figures 2 and 3). In line with these expectations, we observe that more extended nanostructures and spheroids display lower magnitudes of electrical potential, both at the molecular surface (ψ_s) and in the electrolyte (ψ), and are associated with larger magnitudes of effective charge (Figure 3). Comparing calculations of effective charge with the measured values for DNA nanostructures as a function of R_g reveals that the overall trend in measured effective charge, q_m, is captured well by q_{eff} values calculated for the equivalent spheroids (Figure 3). This analysis recovers a feature that is in line with theorems on the geometry dependence of the electrostatic energy and capacity of isoperimetric, charged, conducting solids immersed in a dielectric medium, which have been previously applied to the study of macromolecular topological properties such as knots in polyelectrolytes.[20−24] Our analysis reveals that the most compact charged objects (the sphere and the DNA tetrahedron) have the highest electrical free energy and also the lowest effective charge which may be viewed as analogous to the electrostatic capacity (Figure 3). Our results provide further confirmation that electrometry measurements, in combination with information on a charged molecular species such as a molecular weight estimate, can yield coarse-grained structural information on the 3D morphology of a charged molecule.

Finally, we focus on the 3D conformation of the square-tile. The measurement indicates a value of $q_m = -57.7 \pm 1.5$ *e*, which is significantly larger in magnitude than the nominal value of $q_{eff} = -52.1$ *e* calculated for a planar square-tile structure with an interhelix spacing $s = 2.7$ nm as reported in X-ray scattering measurements.[25,26] We calculated electrostatic free energies in free solution, *F*, and q_{eff} values for various conformations of the square-tile, including a "bent" structure representing the average conformation from 2000 time-steps of a molecular simulation using oxDNA[27,28] (Figure 4a). Under our measurement conditions the planar conformation is suggested to be the thermodynamically favored state as it has the lowest calculated value of *F* (Figure 4a). In general, we expect greater conformational compactness to yield larger electrostatic free energies, corresponding to smaller free energy differences in the trap and therefore smaller magnitudes of q_{eff}. Therefore, an experimental measurement of q_m larger in magnitude than the q_{eff} value of a reference conformational state would *ipso facto* point to a more expanded molecular conformation. Varying the interhelix separation, *s*, of the planar conformation of the tile in calculations, we find agreement between our measured q_m value and the calculated q_{eff} value for

https://doi.org/10.1021/acs.nanolett.2c02485
Nano Lett. 2022, 22, 7834−7840

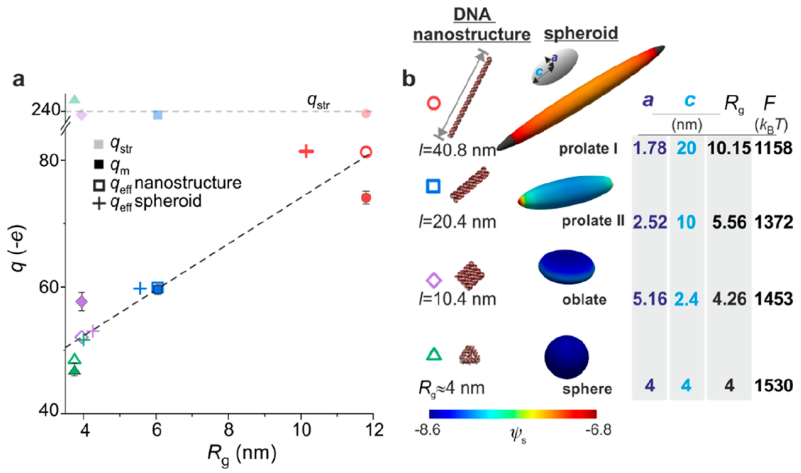


Figure 3. Interpreting the 3D conformation dependence of nanostructure effective charge. (a) Measured q_m (dark solid symbols) and calculated q_{eff} (open symbols) effective charge values for each DNA nanostructure species increase in magnitude with increasing radius of gyration R_g. Dashed lines are guides for the eye. (b) The trend observed for DNA nanostructures is captured by charged hollow spheroids of fixed total volume, $V = \frac{4\pi R^3}{3}$, where R = 4 nm and q_{str} = −240 e. q_m and q_{eff} values for the nanostructures are very similar to q_{eff} values for spheroidal surface charge distributions of similar R_g ("plus" symbols in (a)). We use a sphere of radius R = 4 nm to model the DNA tetrahedron, prolate spheroids I and II to mimic the rod and the bundle respectively, and an oblate spheroid to model the square-tile. In each case we set the length of the longer semiaxis to correspond to half the longest dimension of the corresponding nanostructure. The table lists semiaxis lengths, a and c, for the spheroids, as well as electrical free energies, F, calculated using eq S14 in Supporting Information for each spheroid in free solution (right). F displays an increase with increasing sphericity, whereas q_{eff} decreases with increasing sphericity. Surface electrical potential, ψ_s, distributions display a systematic reduction in magnitude with increasing R_g of the spheroid, correlating with an increase in magnitude of q_{eff}. Note that because all the ionized groups reside on the molecular surface, we use R_g values for hollow rather than solid geometries for the spheroids as well as for the coarse-grained molecular models in Figure 2.

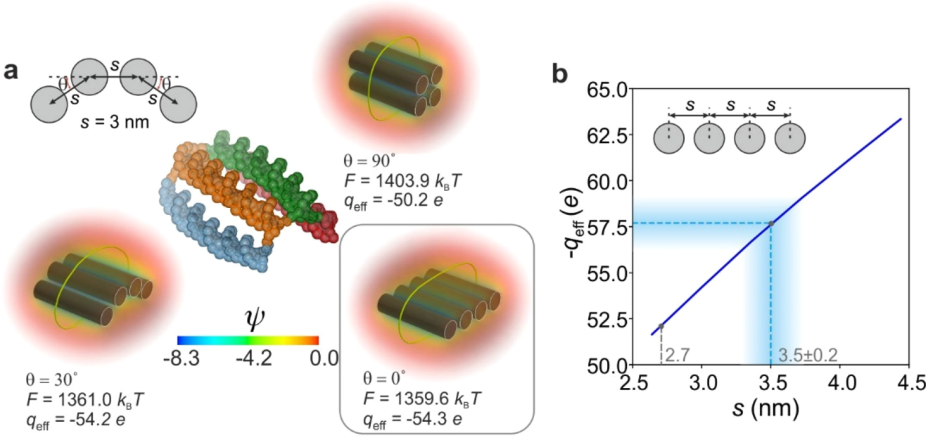


Figure 4. Interhelix spacing in the square-tile nanostructure inferred from its measured effective charge. (a) Calculated electrostatic free energies, F, in free solution (calculated using Supporting Information eq S14) and associated effective charge values, q_{eff}, are shown for three different tile geometries, each characterized by a spacing s = 3 nm and the quoted angle θ for nonplanar structures. The planar structure with θ = 0° corresponds to the state with the minimum electrostatic free energy. Electrical potential distributions, ψ, depict the electrolyte volume surrounding each nanostructure. (b) Plot of calculated q_{eff} vs s for a planar square-tile (θ = 0°). Our measured effective charge, q_m = −57.7 ± 1.5 e, points to a relatively expanded square-tile structure with an interhelical spacing of s = 3.5 ± 0.2 nm.

a structure with s = 3.5 ± 0.2 nm, about 30% larger than the literature value (Figure 4b). This may be explained by the fact that our measurements are performed in solutions containing ∼1−6 mM NaCl where weaker electrostatic screening could be expected to favor a more expanded molecular state. In general, our measurements of the renormalization factor, η, are in good agreement both with our calculated values and with prior theoretical estimates for charged rods and disks (Supporting Information section 3.4).[29,30] Thus, our study further suggests that beyond detecting coarse-grained morphological features of molecular 3D conformation (e.g., elongated/unfolded or compact/globular), precise measurements of electrostatic energies may further shed light on finer-grained structural attributes such as periodicities and spacings in ordered molecular charge distributions.[31]

Our interaction-based definition of molecular effective electrical charge implies that the free energy of interaction between a molecule and a charged entity in solution may be written as $F(\mathbf{r}) = q_{eff}\phi(\mathbf{r})$, where $\phi(\mathbf{r})$ is the electrical potential created by the charged entity at the location of the molecule, $\mathbf{r}$.[8] This definition of interaction free energy is the same as the potential of mean force between a charged object generating a

https://doi.org/10.1021/acs.nanolett.2c02485
Nano Lett. 2022, 22, 7834−7840

potential distribution $\phi(\mathbf{r})$ in the electrolyte and a macroion of effective charge q_{eff} located at $\mathbf{r}$.[13,14] Since the corresponding force can be written as $f(\mathbf{r}) = -q_{\text{eff}}\nabla\phi(\mathbf{r})$, measurements of effective charge can yield a measure of the expected force in the interaction of a molecular species of interest with another charged object in solution. Although our measurements concern the interaction of a molecular scale entity and a like-charged plate, the theoretical basis of the underlying interaction permits us to extrapolate the observed shape-dependence to the interaction between charged molecules in solution.[13,14] The short-range consequences of sequence specificity and precise local structural detail in nucleic acid−protein and protein−protein interactions are well established.[32−34] We have demonstrated that folding, compaction, or conformational collapse of a charged object can reduce the magnitude of the electrostatic energy or interparticle force in an interaction, on account of the effect of *molecular 3D shape alone* on the long-range electrostatics. Macromolecular folding or compaction would normally entail additional energetic effects including contributions from charge regulation or the shifting of acid−base equilibria of ionizable groups and the formation of and exclusion of water from a low dielectric molecular interior,[35−37] all of which might be expected to reduce the structural charge of the molecule by lowering the extent of ionization of its chargeable sites.[38] However, our study utilizes molecular species with highly acidic ionizable groups that retain a large degree of exposure to the solvent, enabling us to probe the impact of molecular 3D conformation alone on long-range electrostatic interactions. Our observations indicate that in an interaction with a charged entity, oligonucleotides with the same number of bases but in different states of compaction experience electrostatic interactions that are significantly different, suggesting a functional role of secondary and tertiary structure in regulating longer-range intermolecular interactions. In other words, molecular shape matters even at long range. We point out that our observations pertain to a Poisson−Boltzmann mean field description of the electrostatic part of an interaction alone, which acts in concert with a number of additional mechanisms, e.g., dispersion, solvation, and fluctuation forces, depending on experimental conditions.[37−41] Finally, our ability to detect differences in the 3D distribution of charge in a molecule is in line with expectations from a rigorous theoretical treatment of multipolar charge distributions in electrolytes.[2,3,14]

In conclusion, the ability to experimentally detect differences in 3D conformation has immediate and obvious implications for the rapid, solution-phase characterization of biomolecular structure and conformation, particularly for strongly charged nucleic acids. For example, we envision the ability to assess the conformations of complex RNA molecules, where structural studies lag far behind those for proteins.[42] The experiments performed in this study were designed to furnish a proof of the principle that precise measurement of electrostatic interactions can yield information on 3D macromolecular conformation in solution. Although ensemble-averaging in data analysis means that measurements reflect the statistically dominant conformation, we expect that in future, analyzing escape time data on the basis of migration trajectories of individual molecules could provide a view of the "conformation spectrum" of a mixture of chemically similar yet structurally and/or conformationally distinct molecular states.

■ ASSOCIATED CONTENT

Supporting Information

Supporting Information (PDF): . The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.nanolett.2c02485.

(1) DNA nanostructure synthesis and characterization; (2) escape-time electrometry (ET*e*) to measure the effective charge, q_m, of DNA nanostructures; (3) procedure to calculate values of effective charge, q_{eff}, for DNA nanostructures (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Madhavi Krishnan − *Physical and Theoretical Chemistry Laboratory, Department of Chemistry, University of Oxford, Oxford OX1 3QZ, United Kingdom; The Kavli Institute for Nanoscience Discovery, Oxford OX1 3QU, United Kingdom;* orcid.org/0000-0003-1274-7155; Email: madhavi.krishnan@chem.ox.ac.uk

Authors

Gunnar Kloes − *Physical and Theoretical Chemistry Laboratory, Department of Chemistry, University of Oxford, Oxford OX1 3QZ, United Kingdom;* orcid.org/0000-0002-7112-8621

Timothy J. D. Bennett − *Physical and Theoretical Chemistry Laboratory, Department of Chemistry, University of Oxford, Oxford OX1 3QZ, United Kingdom;* orcid.org/0000-0001-7419-7756

Alma Chapet-Batlle − *Clarendon Laboratory, Department of Physics, University of Oxford, Oxford OX1 3PU, United Kingdom*

Ali Behjatian − *Physical and Theoretical Chemistry Laboratory, Department of Chemistry, University of Oxford, Oxford OX1 3QZ, United Kingdom*

Andrew J. Turberfield − *Clarendon Laboratory, Department of Physics, University of Oxford, Oxford OX1 3PU, United Kingdom; The Kavli Institute for Nanoscience Discovery, Oxford OX1 3QU, United Kingdom;* orcid.org/0000-0002-3876-0190

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.nanolett.2c02485

Author Contributions

§G.K. and T.J.D.B. contributed equally to the work. G.K. and T.J.D.B. performed experiments and analyzed the data. A.C.-B. and A.J.T. synthesized and provided expertise on DNA nanostructures. A.B. performed molecular modeling. M.K. designed the project and wrote the manuscript. All authors contributed to the manuscript.

Notes

The authors declare no competing financial interest.
Raw data and analysis scripts are available upon request.

■ ACKNOWLEDGMENTS

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 Research and Innovation Programme (Grant Agreement 724180).

https://doi.org/10.1021/acs.nanolett.2c02485
Nano Lett. 2022, 22, 7834−7840

REFERENCES

(1) Manning, G. S. Limiting laws and counterion condensation in polyelectrolyte solutions I. colligative properties. *J. Chem. Phys.* **1969**, *51* (3), 924−933.
(2) Ramirez, R.; Kjellander, R. Dressed molecule theory for liquids and solutions: An exact charge renormalization formalism for molecules with arbitrary charge distributions. *J. Chem. Phys.* **2003**, *119* (21), 11380−11395.
(3) Ramirez, R.; Kjellander, R. Effective multipoles and Yukawa electrostatics in dressed molecule theory. *J. Chem. Phys.* **2006**, *125* (11), 144110.
(4) Manning, G. S. Electrostatic free energies of spheres, cylinders, and planes in counterion condensation theory with some applications. *Macromolecules* **2007**, *40* (22), 8071−8081.
(5) Netz, R. R.; Orland, H. Variational charge renormalization in charged systems. *Eur. Phys. J. E Soft Matter* **2003**, *11* (3), 301−311.
(6) Aubouy, M.; Trizac, E.; Bocquet, L. Effective charge versus bare charge: an analytical estimate for colloids in the infinite dilution limit. *J. Phys. A* **2003**, *36* (22), 5835−5840.
(7) Alexander, S.; et al. Charge renormalization, osmotic pressure, and bulk modulus of colloidal crystals: Theory. *J. Chem. Phys.* **1984**, *80* (11), 5776−5781.
(8) Krishnan, M. A simple model for electrical charge in globular macromolecules and linear polyelectrolytes in solution. *J. Chem. Phys.* **2017**, *146* (20), 205101.
(9) Belloni, L. Ionic condensation and charge renormalization in colloidal suspensions. *Colloids and Surfaces A: Physiochemical and Engineering Aspects* **1998**, *140* (1), 227−243.
(10) Ruggeri, F.; et al. Single-molecule electrometry. *Nat. Nanotechnol.* **2017**, *12* (5), 488−495.
(11) Stellwagen, E.; Stellwagen, N. C. Electrophoretic mobility of DNA in solutions of high ionic strength. *Biophys. J.* **2020**, *118* (11), 2783−2789.
(12) Manning, G. S. Limiting laws and counterion condensation in polyelectrolyte solutions 7. electrophoretic mobility and conductance. *J. Phys. Chem.* **1981**, *85* (11), 1506−1515.
(13) Kjellander, R.; Mitchell, D. J. Dressed ion theory for electric double layer structure and interactions; an exact analysis. *Mol. Phys.* **1997**, *91* (2), 173−188.
(14) Ulander, J.; Greberg, H.; Kjellander, R. Primary and secondary effective charges for electrical double layer systems with asymmetric electrolytes. *J. Chem. Phys.* **2001**, *115* (15), 7144−7160.
(15) Rothemund, P. W. K. Folding DNA to create nanoscale shapes and patterns. *Nature* **2006**, *440* (7082), 297−302.
(16) Goodman, R. P.; Schaap, I. A. T.; Tardin, C. F.; Erben, C. M.; Berry, R. M.; Schmidt, C. F.; Turberfield, A. J. Rapid chiral assembly of rigid DNA building blocks for molecular nanofabrication. *Science* **2005**, *310* (5754), 1661−1665.
(17) Ruggeri, F.; Krishnan, M. Spectrally resolved single-molecule electrometry. *J. Chem. Phys.* **2018**, *148* (12), 123307.
(18) Ruggeri, F.; Krishnan, M. Entropic trapping of a singly charged molecule in solution. *Nano Lett.* **2018**, *18* (6), 3773−3779.
(19) Ruggeri, F.; Krishnan, M. Lattice diffusion of a single molecule in solution. *Phys. Rev. E* **2017**, *96* (6), 062406.
(20) Pólya, G. Estimating electrostatic capacity. *Am. Math. Mon.* **1947**, *54* (4), 201−206.
(21) Pólya, G.; Szegö, G. Definitions, Methods and Results. In *Isoperimetric Inequalities in Mathematical Physics*; Princeton University Press, 1951; Chapter I, pp 1−41.
(22) Vargas-Lara, F.; Mansfield, M. L.; Douglas, J. F. Universal interrelation between measures of particle and polymer size. *J. Chem. Phys.* **2017**, *147* (1), 014903.
(23) Mansfield, M. L.; Douglas, J. F.; Garboczi, E. J. Intrinsic viscosity and the electrical polarizability of arbitrarily shaped objects. *Phys. Rev. E* **2001**, *64* (6), 061401.
(24) Vargas-Lara, F.; Hassan, A. M.; Mansfield, M. L.; Douglas, J. F. Knot Energy, Complexity, and Mobility of Knotted Polymers. *Sci. Rep.* **2017**, *7* (1), 13374.
(25) Fischer, S.; et al. Shape and interhelical spacing of DNA origami nanostructures studied by small-angle x-ray scattering. *Nano Lett.* **2016**, *16* (7), 4282−4287.
(26) Baker, M. A. B.; et al. Dimensions and Global Twist of Single-Layer DNA Origami Measured by Small-Angle X-ray Scattering. *ACS Nano* **2018**, *12* (6), 5791−5799.
(27) Ouldridge, T. E.; Louis, A. A.; Doye, J. P. K. Structural, mechanical, and thermodynamic properties of a coarse-grained DNA model. *J. Chem. Phys.* **2011**, *134* (8), 085101.
(28) Poppleton, E.; Romero, R.; Mallya, A.; Rovigatti, L.; Sulc, P. OxDNA.org: a public webserver for coarse-grained simulations of DNA and RNA nanostructures. *Nucleic Acids Res.* **2021**, *49* (W1), W491−W498.
(29) Agra, R.; Trizac, E.; Bocquet, L. The interplay between screening properties and colloid anisotropy: Towards a reliable pair potential for disc-like charged particles. *Eur. Phys. J. E* **2004**, *15* (4), 345−357.
(30) Téllez, G.; Trizac, E. Exact asymptotic expansions for the cylindrical Poisson−Boltzmann equation. *J. Stat. Mech.: Theory Exp.* **2006**, *2006* (06), P06018−P06018.
(31) Bespalova, M.; Behjatian, A.; Karedla, N.; Walker-Gibbons, R.; Krishnan, M. Opto-Electrostatic Determination of Nucleic Acid Double-Helix Dimensions and the Structure of the Molecule−Solvent Interface. *Macromolecules* **2022**, *55* (14), 6200−6210.
(32) Seeman, N. C.; Rosenberg, J. M.; Rich, A. Sequence-specific recognition of double helical nucleic-acids by proteins. *Proc. Natl. Acad. Sci. U.S.A.* **1976**, *73* (3), 804−808.
(33) Theodoor, J.; Overbeek, G. The role of energy and entropy in the electrical double layer. *Colloids Surf.* **1990**, *51*, 61−75.
(34) Honig, B.; Nicholls, A. Classical electrostatics in biology and chemistry. *Science* **1995**, *268* (5214), 1144−1149.
(35) Sharp, K. A.; Honig, B. Electrostatic interactions in macromolecules - theory and applications. *Annu. Rev. Biophys. Biophys. Chem.* **1990**, *19*, 301−332.
(36) Ninham, B. W.; Parsegian, V. A. Electrostatic potential between surfaces bearing ionizable groups in ionic equilibrium with physiologic saline solution. *J. Theor. Biol.* **1971**, *31* (3), 405−428.
(37) Isom, D. G.; Castañeda, C. A.; Cannon, B. R.; García-Moreno E., B. Large shifts in pKa values of lysine residues buried inside a protein. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108* (13), 5260−5265.
(38) Isom, D. G.; Castañeda, C. A.; Cannon, B. R.; Velu, P. D.; García-Moreno E., B. Charges in the hydrophobic interior of proteins. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107* (37), 16096−16100.
(39) Kirkwood, J. G.; Shumaker, J. B. Forces between protein molecules in solution arising from fluctuations in proton charge and configuration. *Proc. Natl. Acad. Sci. U.S.A.* **1952**, *38* (10), 863−871.
(40) Kubincová, A.; Hünenberger, P. H.; Krishnan, M. Interfacial solvation can explain attraction between like-charged objects in aqueous solution. *J. Chem. Phys.* **2020**, *152* (10), 104713.
(41) Chremos, A.; Douglas, J. F. Communication: Counter-ion solvation and anomalous low-angle scattering in salt-free polyelectrolyte solutions. *J. Chem. Phys.* **2017**, *147* (24), 241103.
(42) Townshend, R. J. L.; et al. Geometric deep learning of RNA structure. *Science* **2021**, *373* (6558), 1047−1051.

https://doi.org/10.1021/acs.nanolett.2c02485
Nano Lett. 2022, 22, 7834−7840

Statement of Authorship for joint/multi-authored papers for PGR thesis

To appear at the end of each thesis chapter submitted as an article/paper

The statement shall describe the candidate's and co-authors' independent research contributions in the thesis publications. For each publication there should exist a complete statement that is to be filled out and signed by the candidate and supervisor **(only required where there isn't already a statement of contribution within the paper itself).**

Title of Paper	Far-field electrostatic signatures of macromolecular 3D conformation
Publication Status	x Published □ Accepted for Publication □Submitted for Publication □Unpublished and unsubmitted work written in a manuscript style
Publication Details	Gunnar Kloes†,Timothy J. D. Bennett†, Alma Chapet-Batlle, Ali Behjatian, Andrew J. Turberfield, and Madhavi Krishnan *Nano Letters* **2022** *22* (19), 7834-7840, DOI: 10.1021/acs.nanolett.2c02485 †authors contributed equally

Student Confirmation

Student Name:	Timothy Bennett		
Contribution to the Paper	Performed experiments and analysed data. Contributed to the writing of the manuscript and supporting information.		
Signature	TBennett	Date	20/9/24

Supervisor Confirmation

By signing the Statement of Authorship, you are certifying that the candidate made a substantial contribution to the publication, and that the description described above is accurate.

Supervisor name and title: Professor Madhavi Krishnan			
Supervisor comments			
Signature	MK-C	Date	20/09/2024

This completed form should be included in the thesis, at the end of the relevant chapter.

2.2 Measurements of post-translational modifications (PTMs)

2.2.1 Why measure PTMs?

Post-translational modifications (PTMs) and other charge modifying reactions are essential to the operation of cells, Sulfonation, for example, improves solubility of dopamine and other endogenous substrates as part of a detoxification pathway in chemical metabolism. PTMs regulate the function of proteins within cells through control of conformation, activity and charge. S-glutathionylation, for example, plays a crucial role in cell signalling, triggering apopotic and proliferative pathways in cells under oxidative or nitrosative stress[76,77]. The importance of these modifications is most evident when dysregulation of PTMs occur, which can cause catastrophic breakdown of normal cellular processes resulting in cancerous growth and disease state formation. For instance, hyperphosphorylation of tau proteins causes aggregation similar to that found in Alzheimer's brain disease[78,79].

2.2.2 PTM detection in thioredoxin

Lots of PTMs alter the structural charge of the underlying protein, as a result ETe is a good candidate method for the detection of such modifications. To investigate the sensitivity of ETe to such modifications, a thioredoxin (TRX) mutant was synthesised and multiple different modifications, including one ATTO532 fluorescent label, were attached and measured. The TRX mutant was synthesised with a 35 amino-acid long linker attached, which included an enzyme recognition site and cysteine for modification as shown in Fig. 5. All modifications were made to this region of the linker. The PTMs were by design charge altering compared to the unmodified protein (WT). The phosphorylated sample (P) had a phosphate group conjugated to the serine, and was expected to give a net -2 e alteration to the q_{str} at pH $>$ 10. The sulfonated sample (S) had a sulfate group conjugated to the cysteine, and was expected to give a net -1 e at pH $>$ -3. The glutathionylated sample (GSH) had a glutathione conjugated to the cysteine, and was expected to give a net -2 e at pH $>$10 . A breakdown of the charged groups and their expected pK_{a} values is shown in table 2.1, although we acknowledge that the specific pK_{a} of a group can depend strongly on the local environment[80]. The samples were measured using ETe, with a 5 nucleotide single-stranded DNA (5ss) labelled with ATTO532 as a calibrator molecule (chosen to be similar in charge to the TRX). The calibrator molecule was measured under the same conditions as the TRX and the surface potential was calculated using Eq. 10, using a charge of 5ss of -6.4e as determined previously[53].

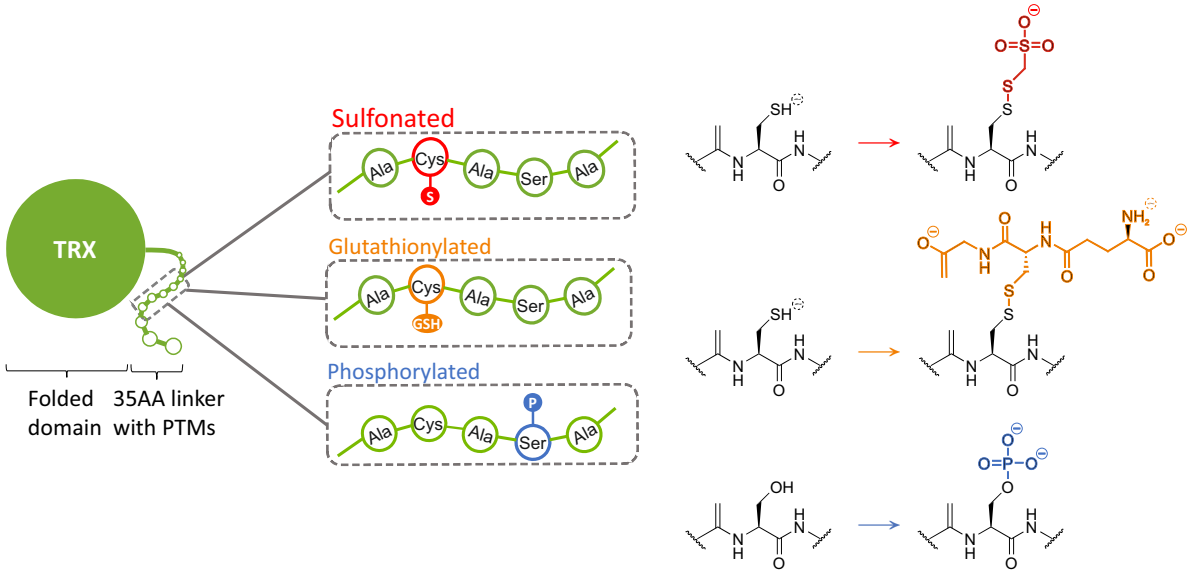


Fig. 5 **Summary of PTMs attached to TRX mutants.** Left panel shows the folded domain of the TRX mutant with 35 amino acid linker attached which includes an enzyme recognition site and cysteine for modifications. Right panel shows the chemical modifications of the PTMs and their approximate ionisation state at pH 9

Table 2.1 **Summary of the relevant charged groups in the modification site on the linker of TRX**[81–84]

Sample	Position (Conjugate Base) pK_a		
Unmodified	Cys (SH) 7.5-9		
Phosphorylated	Cys (SH) 7.5-9	Ser (PO_3H_2) 2	Ser (PO_2H^-) 5.8-7.2
Sulfonated	Cys (SO_3H) -3		
Gulathionylated	Cys (COOH,1) 2	Cys (COOH,2) 4	Cys (NH_2) 8.5-9

The four TRX samples were measured at ca. 1 nM concentration under similar salt concentrations (c=0.26-0.4 mM NaCl effective) at pH 9 in the same device. As shown in Fig. 6, a clear difference in escape time and effective charge between the samples was measured. Note that experimentally measured q_{eff} values might be expected to differ from q_{str} estimates due to effects of both charge regulation and normalisation, see section 1.3. Relative to the unmodified sample, we can rationalise the ca. 1 e higher magnitude charge of the sulfonated version by assuming the cysteine thiol on the unmodified sample is not dissociated at pH 9, this would suggest a slightly more acidic local environment. Similarly we can attribute the ca. 2.5 e higher magnitude charge of the phosphorylated version compared to the unmodified sample to the full dissociation of the phosphate group, and a slightly more dissociated cysteine thiol on the phosphorylated version. However, the ca. 0.5 e higher magnitude charge of the glutathionylated version compared to the unmodified version is much lower than expected, and suggests that neither carboxyllic acid nor the amine in the GSH modification was fully dissociated as we might expect. A closer inspection of the escape time histogram for this sample suggests that this sample is not completely pure, since a bi-exponential fit best captured the underyling distribution, with the longer timescale giving the most realistic charge presented here, see Fig. 6**a**. Some contribution from unconjugated dye, or unmodified TRX could contribute to this heterogeneity. We also note that the escape time histogram measured for the unmodified sample was best fitted with a bi-exponential distribution, whose smaller timescale seems to capture the effective charge expected, while the longer timescale perhaps captures the dynamics of some aggregate species. The escape time histograms for all the species are show in Fig. 6 (**a**).

2.2.3 PTM identification in thioredoxin

In general, not all PTMs are expected to give distinct charge modifications at a specific measurement pH. In order to positively identify a given PTM, a characteristic charge signature over a pH range, governed by the pK_{a} of the charge groups and local environment, is required. Even in the case of a PTM which does not introduce a charged group, one might expect the modification to alter the local environment so as to slightly shift the dissociation of other groups in the vicinity. To test the sensitivity of ETe to a measurement of pK_{a}, the TRX samples were measured at a series of different pH values.

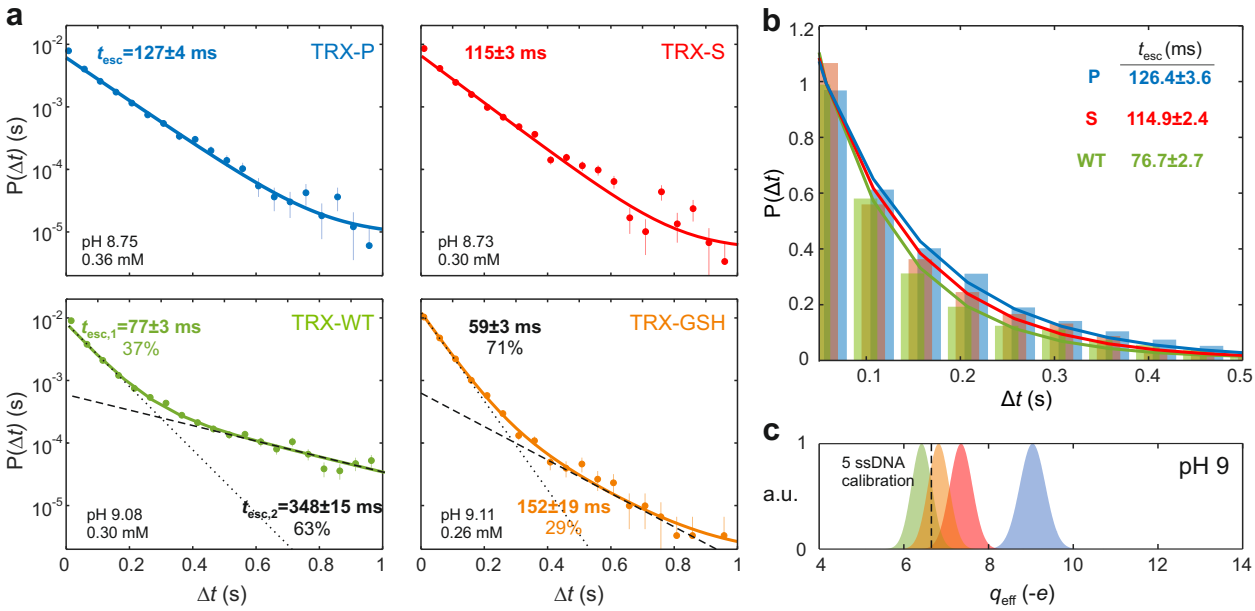


Fig. 6 **Summary of ETe measurements of four TRX samples at pH 9** (**a**) shows the escape time histograms of measurements at pH 9. Lower panels are both fitted with a bi-exponential distribution and show the heterogeneity of the glutathionylated and unmodifed samples. (**b**) overlays the escape time histograms of unmodified, sulfonated and phosphorylated samples. (**c**) shows the q_{eff} inferred from the escape times assuming $D_{\mathrm{t}} = 19\mathrm{x}10^{-7}$ cm^2 s^{-1} and using ϕ_{s} as calibrated from measurements of 5ss in the same device ($2h$=74.5 nm, d=290 nm) under similar conditions. q_{eff} uncertainties are calculated as outlined in Appendix E.

Surface treatment to control ϕ_{s}

In order to facilitate ETe measurements at varying pH, the device surfaces must be characterised over the pH range. The SiO_2 surfaces were observed to significantly lose their charge as the pH is reduced to below 8, resulting in negligible sensitivity to q_{eff} at these pH values, see Fig. 7**a**. As a result a heat-treatment procedure involving heating of the device at 400°C for 4 hours was developed; this resulted in a consistent ϕ_{s} down to pH values below 5. To quantify this, 5ss was measured in a heat-treated device over a range of pH (altered using 0-20 mM Tris-HCl, titrated with variable amounts of HCl), and ϕ_{s} extracted, see Fig. 7**a**. A linear fit to ϕ_{s} vs pH was used to interpolate a ϕ_{s} at any measured pH for further measurements of TRX, discussed in the following section.

Silica surface properties are governed by the variety of different silanol environments exposed to the solution. Recent studies have shown that these groups can vary in pK_a from 2-10, and structures have been assigned to three distinct environments, explaining the differences with hydrogen bonding schemes, see Fig. 7**b**[85–87]. The three structures can be linked to the possible exposed silanol sites, which are defined from the number of neighbouring SiO_4 groups attached to the silanol, and from the distance to the nearest silanol group, as follows: isolated (3 neighbouring

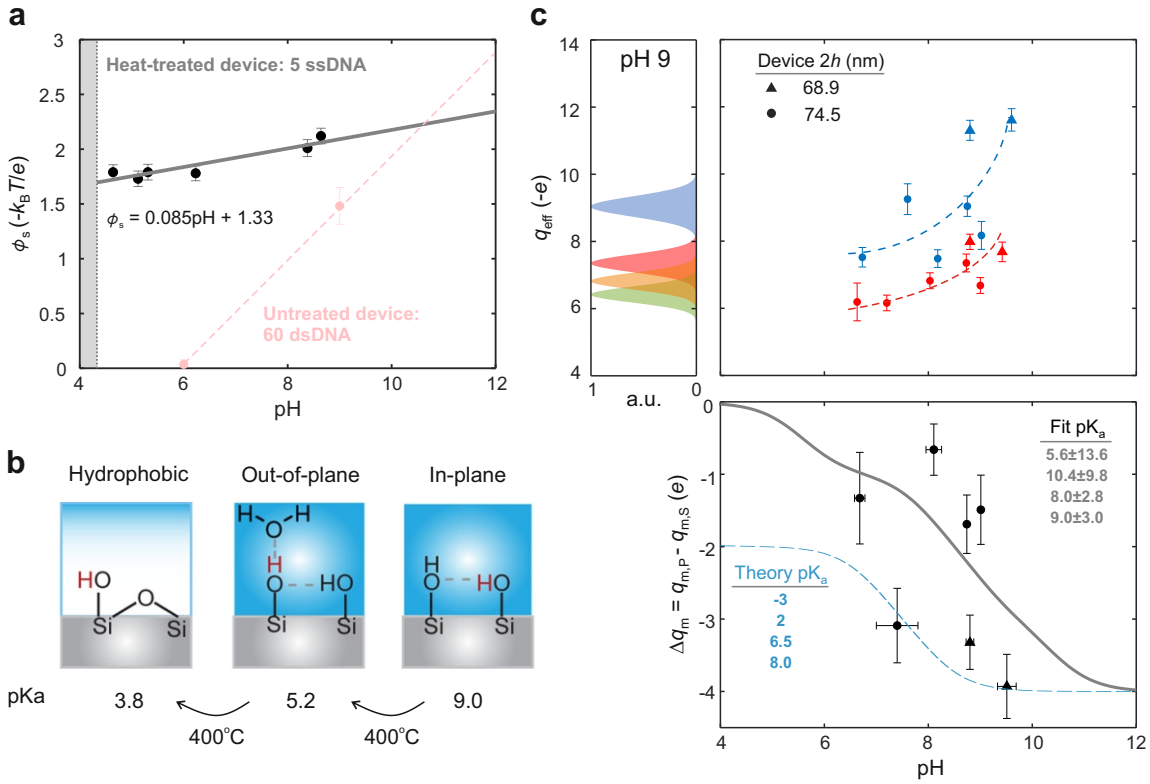


Fig. 7 **Summary of ET*e* measurements of phosphorylated and sulfonated TRX samples at a range of pH** (**a**) shows the calibration curve: a series of measurements of 5ss at different pH values in a heat-treated device, evaluated to extract the implied ϕ_s, and fitted with a linear relation. Overlayed (pink data) is the response in a non-heat-treated device, calibrated using 60 bp DNA, dashed line is a guide to the eye. (**b**) is a scheme adapted from Darlington et al.[85] illustrating the different silanol environments and their respective pK_a values on a water-silica surface. We hypothesise that the heat treatment procedure we use alters the relative populations of these silanols. (**c**) shows the measured q_{eff} for the four TRX samples at pH 9 (left spectra, see Fig. 6). The top right panel shows q_{eff} for phosphorylated and sulfonated samples over a range of pHs values, measured in two heat-treated devices (circles - $2h$=74.5 nm, d=290 nm; triangles - $2h$=68.9 nm, d=290 nm). The lines are a guide to the eye. The bottom panel shows the difference in q_{eff} between the two samples, fitted with a four term pK_a function. q_{eff} uncertainties are calculated as outlined in Appendix E.

SiO_4, no silanol within 3.3 Å (H-bonding range)), vicinal (3 neighbouring SiO_4, nearest silanol is attached to a common O) and geminal (2 neighbouring SiO_4). The low pK_a structure is characterised by an isolated silanol in a locally hydrophobic environment such as an adjacent siloxane, or cyclic trimer[88]. The other two structures are characterised by an out-of-plane H-bond to water and an in-plane H-bond to an adjacent SiOH[89]. Viscinal and geminal silanols are hypothesised to preferentially form one such structure or another[90]. Thus the relative proportion of the three silanol sites could dictate the average pK_a of the surface. Heat treatment of silica surfaces has been shown to affect these relative proportions[91,92]. Here we attribute the increased stability of the device surface charge at low pH upon heat-treatment to the increased acidity of

the silica surface, possibly due to a shift in the relative proportions of the three silanol H-bonding environments.

pK_a measurements for PTM identification

The phosphorylated and sulfonated TRX samples were measured using ET*e* over a range a pH, see Fig. 7**c**. To extract q_{eff}, the surface potential in each device was calibrated using a 5 ss measurement at a single pH, and then using the relation in Fig. 7**a**, extrapolated to find the ϕ_s at the measurement pH. The glutathionylated and unmodified samples were not measured further due to the uncertainty in interpreting their escape time histograms.

The difference in charges of the phosphorylated and sulfonated samples, Δq_{eff}, at each pH was fitted with a curve aiming to extract the four pK_a values expected for the modifications as outlined in Table 2.1, see lower panel in Fig. 7**c**. The fit curve recovers the quantitative result that maximal difference is expected at pH>10 when the cysteine thiol is fully dissociated. Also qualitatively recovered is that at pH≈6, Δq_{eff} is expected to be considerably reduced compared to at pH>10, since neither the PO_2H^- nor the cysteine thiol are dissociated. The exact pK_a values fitted suggest that the first dissociation of the phosphate PO_3H_2 occurs at a slightly higher pH than the literature reports, and is reflected by the smaller magnitude $\Delta q_{\text{eff}} \approx -1$ e at pH≈6, than the $\Delta q_{\text{eff}} \approx -2$ e expected.

2.2.4 Outlook

These measurements stand as a proof of principle for using ET*e* as a PTM detection tool, although we acknowledge some shortcomings in the accuracy in the measurements as presented. The primary difficulty with the measurements presented is that very low salt concentrations (<0.5 mM NaCl) were required. Since the charge of the TRX was very low (ca. -5 e), low salt concentration was required to obtain large ΔF_{el} values to give a high sensitivity to the charge (see Eq. 10 and 29). At these very low salt concentrations, the ionic strengths are hard to control due to effects such as carbon dioxide absorption from air, making accurate measurements difficult. To avoid the requirement of low salt, ΔF_{el} can also be increased by the surface potential and reduction of channel height, evident from Eq. 10. The former is difficult to manipulate due to our constraint of anodic bonding requiring relatively unmodified silica surfaces, whose surface charge was already maximised by measuring at pH>pK_a of all groups, as shown in Fig. 7**a**. The latter was difficult to reduce due to "sagging" of the glass sealing the channels upon anodic bonding[93]. An alternative method of device bonding was recently developed which was immune from the

sagging of the glass lids – thermal bonding. For highest yield of these devices, whole-wafer level bonding was required, as a result a new fabrication workflow was developed and optimised. Full details of this is outlined in Appendix A. Devices with $2h$<20 nm are now routinely possible, which enables larger κ and hence c to be used in an experiment while achieving the same ΔF_{el}. For example, a typical measurement of TRX in a $2h$=70 nm device at c=0.3 mM NaCl gives the same ΔF_{el} as a measurement in a $2h$=20 nm device at c=3.7 mM NaCl - thus more accurate measurements would be possible with the use of these new devices. In fact, such decreases in channel height can facilitate charge measurements in near-physiological salt conditions for highly charged molecules, such as those in section 2.1. In devices with $2h$=20 nm, we expect to be able to use salt concentrations up to 50 mM NaCl and still be sensitive to the charge of the DNA nanostructures.

Measurement of hyperphosphorylated tau protein

To put the PTM detection measurements in a direct biological context, a sample of hyperphosphorylated tau protein was measured using ET*e*. 1 nM of hyperphosphorylated tau was measured in 1.5 mM NaCl 10 nM TCEP (pH 9), and fitted with a tri-exponential distribution, see Fig. 8. The extracted escape times illustrate the broad range of charge states expected in the heterogeneous sample. q_{eff} values were inferred from the escape times and mapped onto the number of phosphate moieties with which the protein could be modified with. Given the sample was measured at pH 9, the phosphates were assumed to be fully dissociated, each contributing -2 e to the charge of the unmodified protein, estimated at +6.3 e at pH 9 using Eq. 2. The inferred phosphorylation states of the samples are in line with mass spectrometry measurements by Meng and coworkers on the same sample[79]. In order to deconvolve the multitude of different charge states in the heterogeneous sample, the single-molecule nature of the ET*e* is required to be exploited – this is explored in the following section 3.

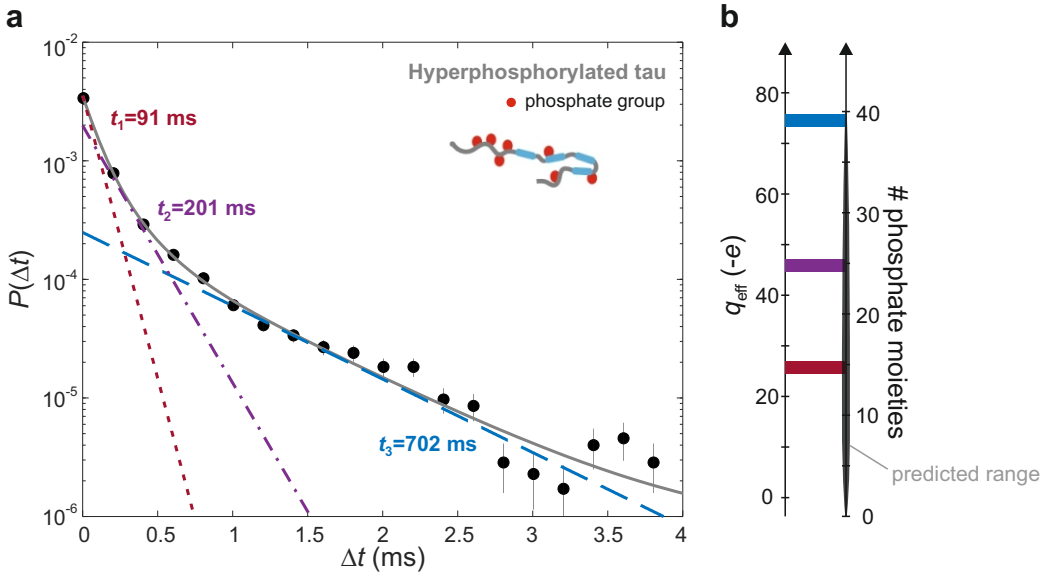


Fig. 8 **Summary of an ET*e* measurements of hyperphorylated tau protein.** (**a**) Escape time histogram of a measurement of hyperphosphporylated tau (schematic from Meng et al.[79]) fitted with a tri-exponential distribution illustrating the range of timescales of species present in the sample. (**b**) q_{eff} values inferred from the timescales fitted in (**a**) assuming $D_{\mathrm{t}} = 5\mathrm{x}10^{-7}\,\mathrm{cm}^2\mathrm{s}^{-1}$ and using $\phi_{\mathrm{s}} = 2.95\ k_{\mathrm{B}}T/e$ as calibrated using a measurement of 60 bp DNA labelled with ATTO532 in the same device ($2h$=75 nm, d=200 nm). The corresponding numbers of phosphate moieties the q_{eff} values imply are also shown, calculated assuming each phosphate contributes -2e to an unmodified protein of charge +6.3e. The measured q_{eff} values match with the range of phosphates expected from MS data reported in the literature (grey shaded region)[79].

Chapter 3

Single-molecule ETe

The motivation for the study of single molecules is clear from section 1.1. This chapter outlines two methods, developed to harness the single molecule nature of ETe to produce spectra distinguishing heterogeneity in complex samples: trajectory-based ETe in section 3.1, and flow-based ETe in section 3.2.

3.1 Trajectory-based ETe (TrETe)

3.1.1 Tracking of single molecules in ETe

Trajectory-based ETe (TrETe), is based on the method of tracking individual molecules in an ensemble ETe experiment[39]; residence times identified to be from a single molecule can be averaged to measure an escape time for that molecule. An effective charge for each molecule measured can then be inferred, allowing single molecule resolution on the charge heterogeneity of a sample.

The experimental design is the same as for ensemble ETe measurements, with a 10-100 fold lower sample concentration used for ease of single particle tracking. Movies were recorded and analysed using the same method as described in section 1.4.3 to extract escape events of length Δt. Single particle trajectories were additionally extracted from the escape events using a custom script; the events are spatially and temporally filtered to try and isolate events originating from a single molecule. Once a molecule trajectory is constructed out of a number, N_{hop}, of escape events, the average event length, t_{av}, is calculated as follows: $\sum_i^{N_{\mathrm{hop}}} \Delta t_i = t_{\mathrm{av}} \approx t_{\mathrm{esc}}$. Histograms of t_{av} from all particle trajectories can be constructed. t_{av} is expected to follow an Erlang distribution with shape factor, $k = N_{\mathrm{hop}}$, and scale factor, $\lambda^{-1} = \frac{\tau}{N_{\mathrm{hop}}}$, where τ is the average escape time of the species of interest. An Erlang distribution is defined as the sum of k independent exponential

variables with mean $\frac{1}{\lambda}$, and is the special case of the Gamma distribution with integer k (see Eq. 16).

$$f(t_{av}, k, \lambda) = \frac{\lambda^k t_{av}^{k-1} e^{-\lambda t_{av}}}{(k-1)!} \tag{16}$$

For $k = N_{hop} \gg 1$, the distribution is well approximated by a normal distribution with mean $\mu = k\lambda^{-1} = \tau$, and standard deviation $\sigma = \sqrt{k}\lambda^{-1} = \frac{\tau}{\sqrt{N_{hop}}}$, so we can interpret the single molecule histograms accordingly. Fig. 9 shows vetting of the expected distribution of t_{av} for varying N_{hop} on simulated data. We clearly see that the number of events collected per molecule limits the precision of a measurement, and consequently the resolving power of heterogeneous states within a sample.

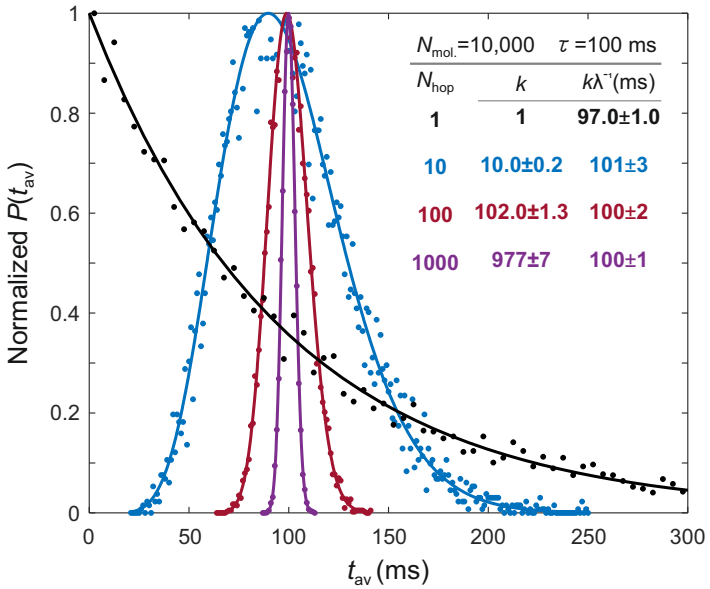


Fig. 9 **Vetting of expected form for t_{av} distribution.** Four histograms of simulated t_{av} distributions from TrETe experiments, each with different numbers of events, N_{hop}, detected per molecule. 10,000 molecule trajectories were simulated in each case and bins of width of 5 ms were used for the histogram. Each distribution was then fitted with an Erlang distribution (Eq. 16) to extract the shape factor, k, and scale factor, λ^{-1}. These values recover the theory that $k = N_{hop}$, and $\lambda^{-1} = \frac{\tau}{N_{hop}}$. We note that the black curve where $N_{hop} = 1$ recovers the exponential distribution typical of ensemble ETe experiments.

3.1.2 Measurements of double-stranded DNA

Pure 60 bp DNA

To first vet the TrETe measurement principle, a pure sample of 60 bp dsDNA was measured in 1.44 mM NaCl (pH 9), see Fig 10. The device was calibrated using the ensemble analysis of the measurement. To verify the dependence of the distribution width on N_{hop}, multiple different

thresholds were set on the number of molecules fitted. As shown in Fig. 10**b**, the fitted k values are close to the minimal value expected of $N_{\mathrm{hop}} = 5$ and $N_{\mathrm{hop}} = 40$ respectively. q_{eff} was calculated for each trajectory and histogrammed similarly. These distributions were fitted with Gaussians, whose mean values, μ, recover the value with which the ensemble data was calibrated with, thus validating the accuracy of the single-molecule realisation.

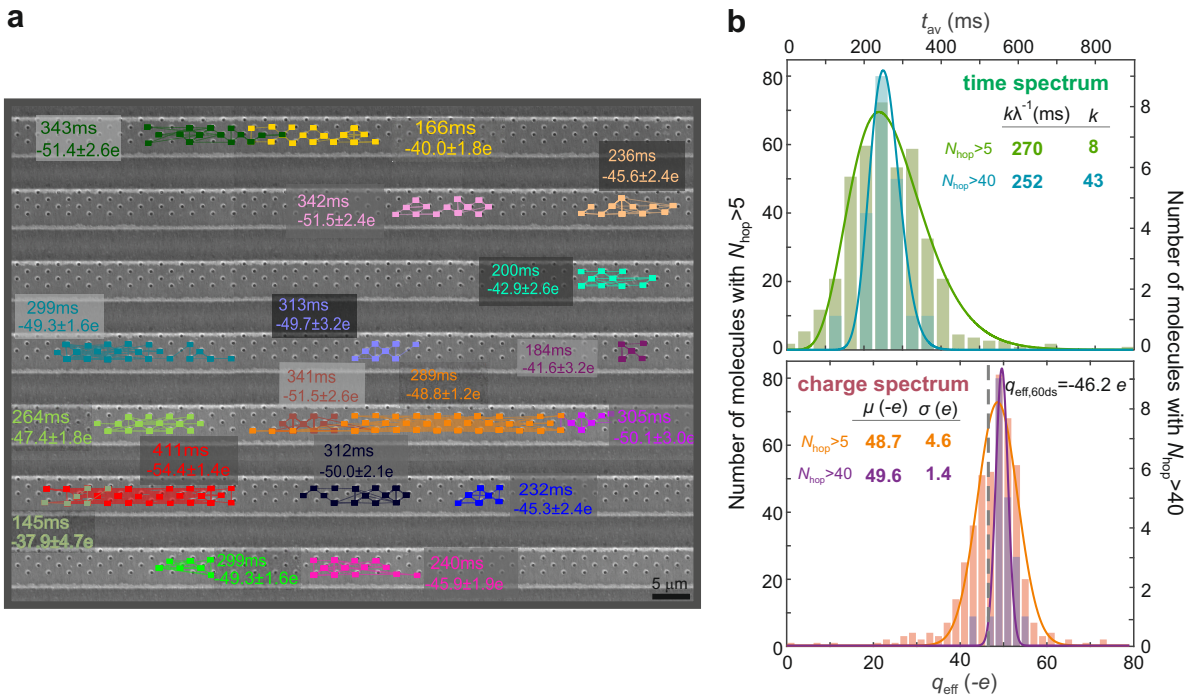


Fig. 10 **TrET*e* measurement of 60 bp DNA.** (**a**) Single-molecule trajectories overlayed on a representative SEM image of the device. t_{av} can be extracted from each molecule, and a q_{eff} calculated. (**b**) Single molecule spectra of t_{av} (upper panel) are produced from trajectories with $N_{\mathrm{hop}} > 5$ (green bars) and $N_{\mathrm{hop}} > 40$ (blue bars). The respective distributions are fit with Erlang distributions (Eq. 16), whose shape parameter, k, matches up closely with the minimum expected value of $N_{\mathrm{hop}} > 5$ and 40 respectively. Single molecule spectra of q_{eff} (lower panel) are calculated for each trajectory assuming $D_{\mathrm{t}} = 5\mathrm{x}10^{-7}\ \mathrm{cm}^2\,\mathrm{s}^{-1}$, based on a device calibration using the ensemble data; $t_{\mathrm{esc}} = 250$ ms for a fixed $q_{\mathrm{eff}} = -46.2\ e$, yields $\phi_{\mathrm{s}} = -2.86\ k_{\mathrm{B}}T/e$ for a device with $2h = 70.9$ nm and $d = 170$: nm. The respective distributions ($N_{\mathrm{hop}} > 5$, orange bars, and $N_{\mathrm{hop}} > 40$, purple bars) are fit with Gaussian distributions with mean μ, and standard deviation σ.

1:1 mixture of 30 and 60 bp DNA

To further vet the system, a 1:1 mixture of 30 bp and 60 bp dsDNA was measured in 1.29 mM NaCl (pH 9), see Fig 11. The device was calibrated using an ensemble analysis of a 60 bp DNA measurement under similar conditions. We can clearly resolve the two states in both the t_{av} (panel (**a**)) and q_{eff} (panel (**b**)) distributions. The t_{av} distribution is fitted with a sum of Erlang distributions (Eq. 16) with fixed $k_i = 20$, corresponding to the minimal value of N_{hop} as guided by the observations in Fig. 10, and we recover the ca. 50% molecule fraction, m_i, expected

from the amplitudes. However we note that the underlying distributions are slightly broader than the theoretical limit, a fit with free k_i yields values of $k_i < 20$; this could be indicative of sample heterogeneity or additional measurement noise, such as channel height variability or mis-assignment of molecular trajectories. Channel height variability can have a large impact on the well depth of a trap since it exponentially affects the electrostatic interaction energy, see Eq. 10. We estimate that a 2 nm variation in $2h$ (3%) results in a 14% uncertainty in the q_{eff} - this is the maximum variation in channel heights we expect over a device. A sum of two gaussians fit to the q_{eff} distribution recovers mean values μ_i that are consistent with the calculated values of 60ds (-46.2 e)[74], and 30ds (-25.8 e)[54] respectively.

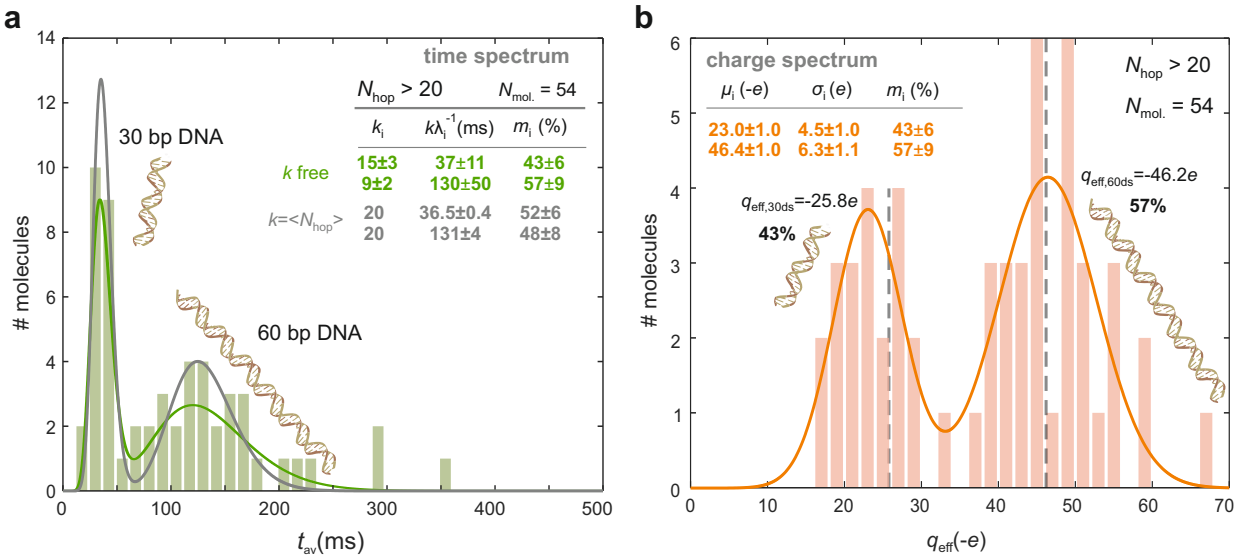


Fig. 11 **TrET*e* measurement of a 1:1 mixture of 30 bp and 60 bp DNA.** (**a**) Single molecule t_{av} spectra were created from trajectories with $N_{\mathrm{hop}} > 20$, and fitted with a sum of two Erlang distributions (Eq. 16). Based on observations in Fig. 10, the distributions were fit with $k_i = 20$ representing the theoretically broadest distribution expected (grey lines), in this instance we recover the ca. 50% molecule fraction, m_i expected. We see that leaving k as a free parameter we see additional broadening of the distributions, since $k_i < 20$ (green lines). (**b**) Single molecule spectra of q_{eff} are calculated for each trajectory assuming $D_{\mathrm{s}} = 7.7\mathrm{x}10^{-7}$ cm^2 s^{-1} and using $\phi_{\mathrm{s}} = -2.70\, k_{\mathrm{B}}T/e$ as calibrated using an ensemble measurement of 60 bp DNA in the same device ($2h$=75 nm and d=163 nm). The distribution is fit with a sum of two Gaussian distributions with means μ_i, that match closely with the calculated values[74,54]. The amplitudes recover the ca. 50% expected.

3.1.3 Outlook

These measurements stand by way of proof of principle of TrET*e*, although we identify possible improvements to the current setup.

Methods to improve the throughput of TrET*e* experiments

The throughput currently is limited by the manual trajectory construction which requires very low concentrations. To improve this, various single-particle tracking algorithms were looked into including TrackNTrace[94] and TrackMate[95] with some success. We expect that the addition of maximum likelihood estimations in the tracking will improve fidelity and enable a higher density of particles to be tracked. Generally we anticipate that the successful implementation of an automated tracking procedure would greatly reduce the analysis time. thus increasing the throughput of the measurement.

The throughput can also trivially be improved by increasing the microscope field of view (FOV), and hence parallelising the measurement. This was improved 20-fold by the microscope improvements outlined in Appendix B, and the impacts on TrET*e* are evident from the single-molecule spectra shown in chapter 4 where trajectories from 100s of molecules are typically accessible.

Methods to improve the resolution of TrET*e* experiments

The precision and hence resolution of TrET*e* measurements is limited by the number of events collected per molecule, N_{hop}. In experiments, two phenomena limit the amount of observation time available for measurement of escape events per molecule: diffusion of molecules out of the microscope field of view and photobleaching of the fluorescent label. The larger field of view microscope previously mentioned improved the former, more details in Appendix B. To combat the photobleaching, the background noise of the microscope was minimised to accommodate the lowest laser power densities possible, while still achieving the signal-to-noise ratio (SNR) required for molecule detection, again discussed in Appendix B. These advancements were used in chapter 4, where the single-molecule experiments regularly yielded $N_{\mathrm{hop}} \gtrsim 100$.

Some chemical solutions to photobleaching were also considered. The photobleaching mechanism for organic fluorophores, as in use here, often involves a stochastic excursion of the dye to a triplet excited state, which is then perceptible to permanent degradation via reactions with reactive oxygen species (ROS) in solution. Such species are often formed by redox reactions of the triplet state of the fluorophore itself with molecular oxygen dissolved in solution, thus reduction of molecular oxygen levels is desireable. To that end, the addition of various oxygen scavenging systems has been shown to improve lifetime before photobleaching by up to an order of magnitude[96]. One such system is a mixture of glucose, glucose oxidase and catalase (GOD:CAT).

One downside to the removal of molecular oxygen from solution is that a key mechanism for return of the fluorophore from the triplet state (which is non emitting) to the ground state, is via redox reactions. Without availability of an oxidising agent for the triplet state fluorophore to react with, it can remain non-emitting for long timescales (ms - hours) - photoblinking. Long periods of silence from fluorophores has impacts on single-molecule tracking, as well as SNR and the escape times extracted, see Appendix F for more details. As a result, various reducing and oxidising systems (ROXS) have been developed to facilitate the rapid oxidation of triplet states without the formation of ROS, and hence stabilise the photoblinking without causing photobleaching. An example of such a stabilising agent is Trolox. We employed a mixture of Trolox and GOD:CAT, in ratios as prescribed by Roy and coworkers[97], in ETe measurements but saw no measurable impact on SNR or event rate (indicative of bleaching rate). It is well reported that the performance of such photo-stabilisation systems vary from fluorophore to fluorophore, and although success has been shown with the ATTO family of dyes with such systems, ATTO532 is not as widely used as 488, 550, 565 and 647N variants[96,98]. Further investigation into such stabilisation methods or chemically engineered fluorophores - self-healing dyes - could prove fruitful for future trETe measurements[99].

Despite these aforementioned improvements in N_{hop} and hence resolution, there will always be an absolute limitation of photobleaching which will limit the ultimate precision possible, and to answer questions such as the phosphorylation state detection posed in section 2.2, a resolution of $\lesssim -1e$ would be required, corresponding to $N_{\text{hop}} \gtrsim 300$. One such solution for this would be label-free tracking of molecules, such as acheived by Špačková and coworkers in similar microfluidic devices[12], although this brings with it limitations on both the FOV and minimum mass of particle detectable. The following section 3.2 outlines our simple fluorescence-based approach, where the need for constant observation of molecules is not required - FlETe.

3.2 Flow-based ET*e* (FlET*e*)

3.2.1 ET*e* measurements without continuous observation

The desire to increase the resolution of TrET*e* beyond the limitation of fluorescence bleaching inspired the creation of flow-based escape-time electrometry, FlET*e*. Average residence times are measured by recording the time taken for a molecule to escape from a known number of traps. Thus, the need for direct measurement of each residence time is not required, and so photobleaching is not a limiting factor. Spectra of average residence times can be created, revealing single-molecule resolution on the heterogeneity of a sample.

The device design is altered for FlET*e* measurements to ensure molecules pass through a known number of traps. Pockets are replaced with an array of rectangular line-traps which span the entire width of a channel (Fig. 12 shows the device design, and Appendix A provides more details). Molecules are driven along the channels by applying a constant pressure difference along the length of the channel, and are observed only in "observation windows" at various positions along the channel. Because the line-traps span the entire width of the channel the molecule is known to have interacted with each trap as it traverses the length of the channel. The time for a molecule to pass from one observation window to the next is recorded, and since the number of line-traps, N_{trap}, between windows in known, the average interaction time (residence time) with each trap is calculated, to obtain t_{av}.

Several considerations are important when calculating t_{av} in FlET*e* experiments. Since we are measuring the time the molecules take to travel a fixed distance down the channel, t_{FlETe}, then this includes a contribution not only from a number, N_{hop}, of stochastic escapes from potential wells, each with average escape time t_{av}, but also from convection speed of the molecule, t_{conv}, as outlined in Eq. 17.

$$t_{\text{FlETe}} = t_{\text{conv}} + N_{\text{hop}} t_{\text{av}}. \tag{17}$$

We note that theoretically $N_{\text{hop}} = N_{\text{trap}}$, however since molecules are also freely diffusing within the flow, N_{hop} comes with some uncertainty as molecules could diffuse against the flow direction and interact with line-traps multiple times. To account for this we consider the Péclet number, Pe, as defined in Eq. 18 as the ratio of diffusion to convection time between line-traps, for flow

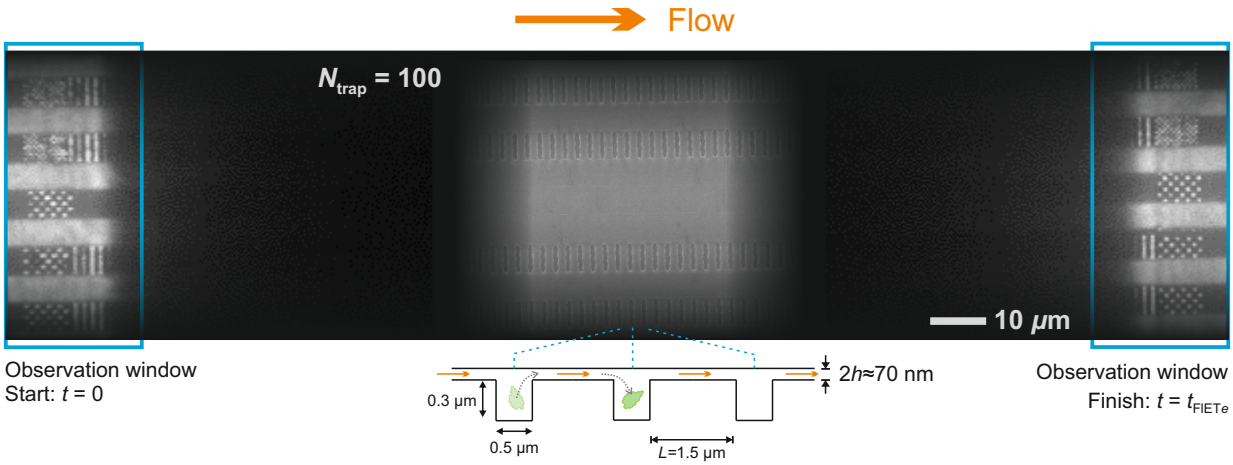


Fig. 12 **Schematic of a FlET*e* measurement.** Left and right extremes show a maximum intensity stack from the fluorescence microscope image observed. Bright horizontal regions are the walls between channels. Other bright regions indicate single molecule signals. Two different pitches and diameters of pockets are displayed in the observation windows. Continuous arrays of N_{trap} line-traps fill the space between observation windows but are obscured from illumination by the excitation mask. Centre of the image shows a representative SEM of a FlET*e* device. Bottom schematic shows the side projection of a single channel: rectangular line-traps ca. 500nm in width, and ca. 300 nm deep are spaced at L≈1.5 μm intervals, edge-to-edge, in an otherwise $2h \approx 70$ nm deep channel. Flow drives molecules from left to right along the channels, the time taken for a molecule to move from the start to finish observation window is measured as t_{FlETe}.

speed u and line-trap edge-to-edge spacing L.

$$\text{Pe} = \frac{Lu}{D_{\text{t}}} \tag{18}$$

Since the flow is pressure driven, both u and hence t_{conv} can be estimated for the experiment by considering the Hagen-Poiseuille equation for a rectangular channel, Eq. 19. The volume flow rate, Q, is evaluated for a pressure drop ΔP, across a channel of length l, filled with fluid of viscosity η, and assuming the channel width w is much larger than the height $2h$. From Q the average velocity, $\langle u \rangle$, and the centreline velocity, $u_{\text{max}} = 1.5\langle u \rangle$, can be calculated.

$$\text{Q} \approx \frac{2\Delta P h^3 w}{3\eta l}\left(1 - 1.26\frac{h}{w}\right) \tag{19}$$

3.2.2 Simulating FlET*e* experiments

To explore the impact different Péclet numbers has on our experiment, a series of BD simulations were setup with variable flow rates. It was found that for typical device geometries when electrostatic interaction energies with the trap $\Delta F_{\text{el}} > 1\ k_{\text{B}}T$, molecular convection speeds down the channel were well approximated by the centre-line velocity, u_{max}. This is due to the

electrostatic repulsion from the channel walls. More details on the BD simulations are included in Appendix D.2.

From the simulations, t_{FlETe} was measured for escape from N_{trap}=100 traps, and using Eq. 17, t_{av} was extracted, see Fig. 13. As in section 3.1.1 we expect t_{av} to follow an Erlang distribution (equation 16), which for N_{hop}=N_{trap}=100 is well approximated by a normal distribution.

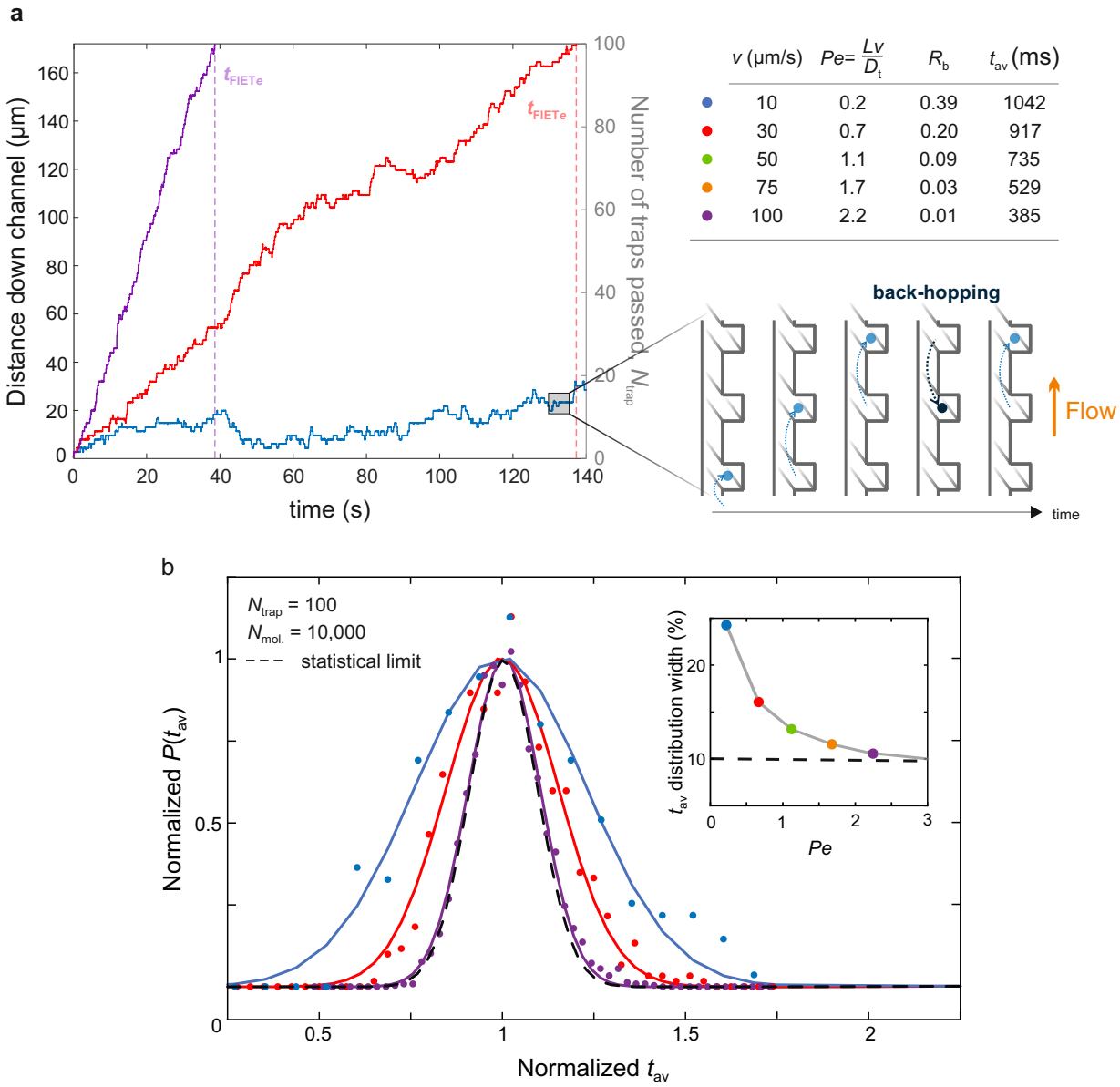


Fig. 13 **Impact of Péclet number on t_{av} distributions in FlET*e* experiments.** (**a**) 1D BD simulations of particles with D_{t} = 5x10^{-7} cm^{2}s^{-1} (as for 60 dsDNA) flowing through a FlET*e* device geometry, with line-traps of width 0.6 μm and spaced edge-to-edge L=1.12 μ m, with different flow speeds imposed in the channels, v=10-100 μm/s. The time taken for a particle to pass N_{hop} = 100 line-traps yields t_{FlETe} times. We see that for slow flow speeds (v=10 μm/s, blue curve) the molecule does not continuously move along the channel with the flow but sometimes "back-hops", lower right panel. (**b**) This back-hopping causes considerable broadening of the t_{av} distributions, calculated using Eq. (17). These distributions are fitted with Gaussians and standard deviations extracted - displayed as a function of Pe in the inset. As $Pe\rightarrow\infty$, we recover the expected statistical limit of $\sigma_{\mathrm{t}} = \frac{t_{\mathrm{av}}}{\sqrt{N_{\mathrm{trap}}}}$.

Diffusional broadening of t_{av} when $Pe < 1$

For $Pe \lesssim 1$ (cases for small L, slow v and/or fast diffusing species) we see diffusional broadening of the normalised t_{av} distribution (Fig. 13**b**); the relatively fast diffusion rate compared to convection rate meaning the molecule interacts with inconsistent numbers of traps, i.e. $N_{\mathrm{hop}} \neq N_{\mathrm{trap}}$. We can see this in panel Fig. 13**a**, which shows the so-called "back-hopping" of the particle against the flow thus meaning it interacts with additional traps. We quantify this with the back-hopping ratio R_{b}, which we expect to be zero as $Pe \to \infty$, and R_{b}=0.5 indicates a particle with no net migration since it on average takes as many forward steps (with the flow), as backward; for the Pe=0.2 case we observe $R_{\mathrm{b}} \approx 0.4$, meaning on average the particle interacts $N_{\mathrm{hop}} \approx$500 times with traps (for every 5 steps forward, the molecule takes 4 steps backward). When we account for this increase in the number of trap interactions, we can evaluate the true t_{av}, displayed in right panel of panel (**a**). We note that this will also include a considerable amount of diffusion time between line-traps (not accounted for by t_{conv} since that assumes no back-hopping). With this in mind, it is hard to deconvolve the differences we observe in t_{av} for the different flow speeds, from being due to back-hopping or being due to fluid velocity perturbations on the escape process; further analysis of simulations is required to determine this. In our experiments we aim for high Péclet numbers where the system is flow limited, to gain the highest certainty in the number of escape events of a molecule, and thus obtain the result statistically limited by N_{trap}.

3.2.3 Optimising FlET*e* experiments

In order to maximise the precision of a measurement, we face a series of engineering challenges to maintain Pe>1 while maximising the number of line-traps between observation windows. For Gen 1 ET*e* devices with $h_1 = 70$ nm, w=5 μm, l=800 μm, $\Delta P = 1$ bar and $\mu = 8.9\text{x}10^{-4}$ Pa s, the centre-line flow velocity in a channel is expected to be $u_{\mathrm{max}} \approx 85$ μm s^{-1} from Eq.19. For 60 bp DNA with $D_{\mathrm{t}} = 5\text{x}10^{-7}$ cm^2 s^{-1}, this constrains L>0.6 μm, whereas for free dye, ATTO532-carboxyl with $D_{\mathrm{t}} = 40\text{x}10^{-7}$ cm^{2}s^{-1}, this constrains L>4.7 μm to maintain Pe>1, which limits the number of line-traps possible between observation windows.

Device design considerations for FlET*e* experiments

From Eq. 18 it is evident that maximising the flow speed maximises Pe for given line-trap separation. Following Eq. 19, the channel length defined by the position of the "step" was minimised to be exactly the width of the observation windows, with l=150 μm we see a 5-fold

increase in Pe, see Appendix A for a detailed schematic of the device. The line-trap width was also reduced to increase the edge-edge spacing of line-traps, L. Cylindrical pockets, as used in ET*e*, of varying sizes were included in the observation zones to improve molecule detection. Devices with variable line-trap spacings were fabricated for use with different molecules.

Channels with no line-traps between observation windows were also included so that t_{conv} could be accurately measured simultaneously and in the vicinity of the $t_{\mathrm{FlET}e}$ measurements. Variation in channel height is possible across the device, and so measurements of flow speeds in neighbouring channels are more likely to be accurate, although we note that height variability is usually less than a few nm (3%), and so flow speeds vary by less than 5%. The variation in height is far more consequential for the electrostatic interaction energy of the traps as discussed in section 3.1.3.

Microscope design for FlET*e* experiments

The upgraded microscope mentioned in section 3.1.3, and outlined in Appendix B was also used for FlET*e* measurements but with a mask in the excitation pathway to create the observation windows. Using this wide-field setup inherently constrains the distance between observation windows to be within the imageable field of view, limiting the maximal N_{trap} possible.

An alternative and cheaper microscope design was also built which used a split excitation beam (using polarising beamsplitting cubes) to illuminate spatially independent parts of the device, meaning a lower power laser and smaller optics could be used.

Data analysis of FlET*e* experiments

Fig. 14 illustrates a typical measurement in a FlET*e* device. Molecules entering and exiting the FOV are identified in the observation windows from high densities of high SNR escape events detected identically to in ensemble ET*e* experiments. Because molecules were not under constant observation, deconvolution of potentially overlapping trajectories was challenging from just start and finish detections. As a result very dilute samples were measured such that there was only a single molecule between windows at one time, and trajectories were then linked manually. The difference in time between detections of the same molecule in the start and finish windows yields $t_{\mathrm{FlET}e}$.

Once $t_{\mathrm{FlET}e}$ and t_{conv} were measured from channels with and without line-traps between observation windows respectively, t_{av} was calculated from Eq. 17 assuming the statistically limited $N_{\mathrm{hop}} = N_{\mathrm{trap}}$ and analysed similarly to TrET*e* data.

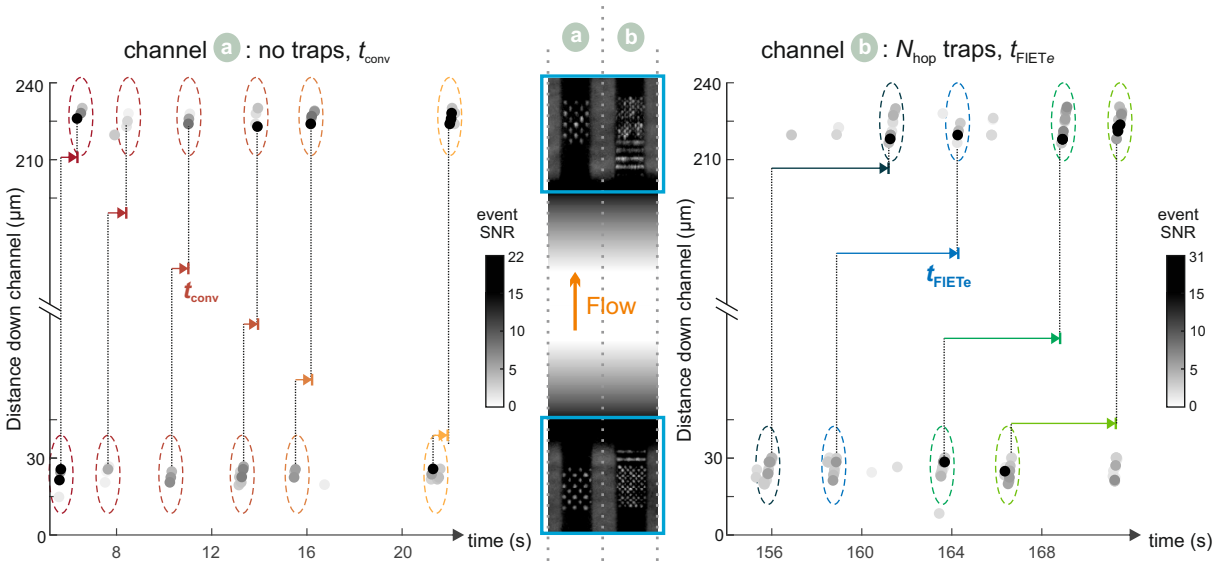


Fig. 14 **FlET*e* measurement principle.** Molecules are detected in an array of pockets in the imaging windows. Escape event locations in space and time for two exemplary channels are shown as grey-scale circles in the left and right panels; SNR is illustrated by grey-scale intensity. Channel (a) has no line-traps, as shown in the maximum intensity fluorescence image in the centre panel, and so is used to measure t_{conv}. Molecules take far longer to migrate between observation windows in channel (b) than (a) due to the interaction with N_{hop} traps, as indicated by $t_{\mathrm{FlET}e} > t_{\mathrm{conv}}$. Data shown is for a sample of 60 bp DNA measured in 1.0 mM NaCl (pH 9) in a $2h$=70 nm device with line-traps of width 560 nm spaced 1500 nm centre-to-centre, pockets in channel (a) are ca. 700 nm diameter, pockets in channel (b) are ca. 500 nm diameter.

3.2.4 Measurements of double-stranded DNA

Pure 60 bp DNA

To first vet the FlET*e* measurement principle, a pure sample of 60 bp dsDNA was measured in 1.3 mM NaCl (pH 9) (see Fig. 15**a**). FlET*e* times were measured between observation windows ca. 230 μm apart, yielding $N_{\mathrm{trap}} = 115$ traps, each 600 nm wide and spaced 2 μm apart centre-to-centre, giving L=1.4 μm. $u_{\mathrm{max}} = 170\,\mu\mathrm{m/s}$ is expected from Eq. 19, giving $Pe{\approx}3$ for this measurement.

t_{av} was inferred using Eq. 17 from measurements of $t_{\mathrm{FlET}e}$. Some channel to channel variability was observed, as shown by the mean t_{av} values for each channel (black boxes in left panel of Fig. 15**a**) varying from the overall distribution mean by more than their respective standard errors. This is reflected in the broader Erlang distribution fitted to the t_{av} histogram than expected (right panel of Fig. 15**a**). Differences such as small variations in channel depth, $2h$, are typical and will affect t_{FlETe} both from variable flow speeds and variable electrostatic interaction strengths, see section 3.2.3. To try and account for this, a normalised t_{av} distribution was generated where each channel has the same mean t_{av} (Fig. 15 **b**). This distribution was

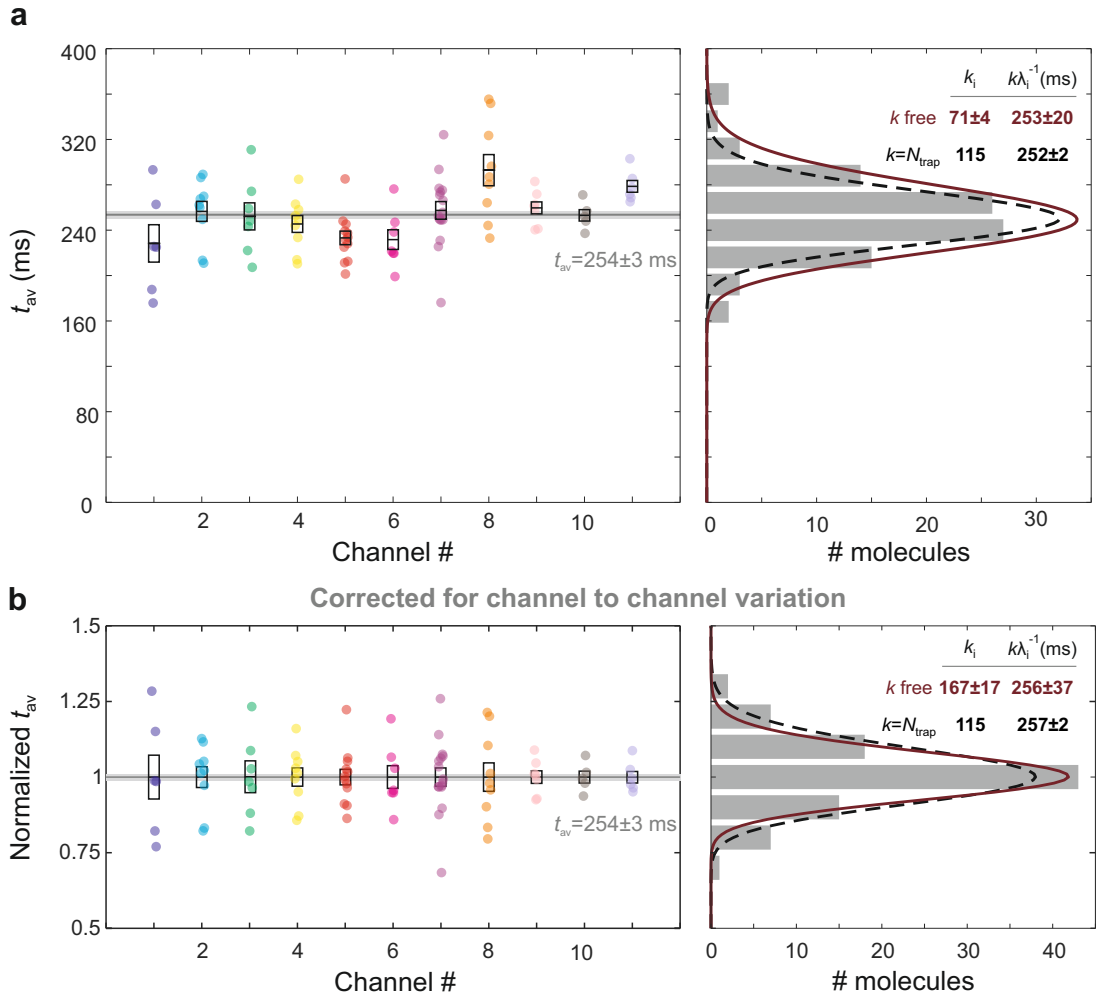


Fig. 15 **Measurement of 60 bp DNA using FlET*e*.** (**a**) t_{av} measurements of 20 pM of 60 bp DNA in a device with $2h$=70 nm, Measurements are displayed per channel to illustrate channel to channel variability (left panel); black boxes illustrate the mean and standard error on the mean of t_{av} for each channel. Distributions of t_{av} are histogrammed with a bin width of 24 ms corresponding to average uncertainty on a single measurement, and fitted with an Erlang distribution (Eq. 16, right panel). A slightly broader distribution than statistically expected was observed (k=71). (**b**) The same distribution was normalised so that each channel has the same average t_{av}, this simulates a correction due to channel to channel variation. A fit to the normalised distribution histogrammed with bin width of 0.1, again representing the average uncertainty on a single value, yields a narrower distribution than statistically expected (k=167).

observed to be narrower than expected after fitting with an Erlang, suggesting that this is not a fair representation of the measurement variation. We acknowledge that more events are required in each channel for this re-normalisation to be fair, and to remove channel to channel variability entirely, only a single channel should be used for measurements. However we also note that k=71 corresponds to a <30% broadening of the distribution compared to the statistical limit of k=115, and so although it is more than is expected given the Péclet number of the experiment (with Pe=3, we expect negligible broadening), the resolution is still useful when it comes to

distinguishing sample heterogeneity. It is also possible that even a nominally pure DNA sample has some inherent heterogeneity that we are measuring here.

1:1 mixture of 30 and 40 bp DNA

To test the resolving capabilities of FlET*e* on a heterogeneous sample, a mixture of 30 and 40 bp DNA samples were measured, as well as the respective pure samples, in 0.9 mM NaCl (pH 9), see Fig. 16. In order to maximise the precision, a device was made with as short a channel as possible to maximise velocity, $l =$ 150 µm giving $u_{\max} \approx 450$ µm s^{-1} from Eq.19; this was validated against a measurement of $t_{\text{conv}} = 320$ ms between observation windows 92 µm apart, yielding u=285 µm s^{-1} in a channel with no line-traps. We attribute the discrepancy to imperfect sealing of the microfluidic device causing some pressure leakage. $t_{\text{FlET}e}$ was measured in channels with line-traps 300 nm in width, spaced 1 µm centre-to-centre apart, giving L=700 nm. For the smallest molecule of interest, 30 bp DNA with $D_{\text{t}} = 7.7\text{x}10^{-7}\text{cm}^2\text{s}^{-1}$, $Pe\approx$3.75 for this measurement, so we expect to have statistically limited t_{av} spectra.

$t_{\text{FlET}e}$ was measured over $N_{\text{trap}} = 92$, t_{av} was calculated using Eq. 17, and single molecule spectra were constructed, shown in Fig. 16. The spectra were fitted with a sum of two Erlang distributions (Eq. 16), and for the pure samples we observe slight broadening of the distribution widths compared with the statistical limit: ca. 18% broader for 30ds and ca. 46% for 40ds. In the mixed sample two states are clearly resolvable, especially when consideration of the expected distribution widths are taken into account. We do note however, a considerable number of molecules with outlying t_{av}. The consistent presence of a peak at much shorter t_{av} could be due to a certain amount of free dye in the samples. The presence of a single molecule with large t_{av} in both spectra of pure species can be attributed either to a molecule which temporarily stuck to the channel surface, or could be due to a molecule which photobleached in the start observation window, and thus exit was only triggered by the next molecule down the channel, for example. These sorts of detection issues were obstacles to clean FlET*e* measurements in general.

3.2.5 Outlook

These measurements show the proof of principle of using FlETe for high precision single molecule measurements, able to capture sample heterogeneity at the escape time level, although we identify lots of room for improvements in the current setup. Further comparison with simulations of different well depths would enable effective charge spectra to be produced, as mentioned in Section 3.2.2, with higher resolving capabilities possible due to the exponential sensitivity of t_{av} on q_{eff}.

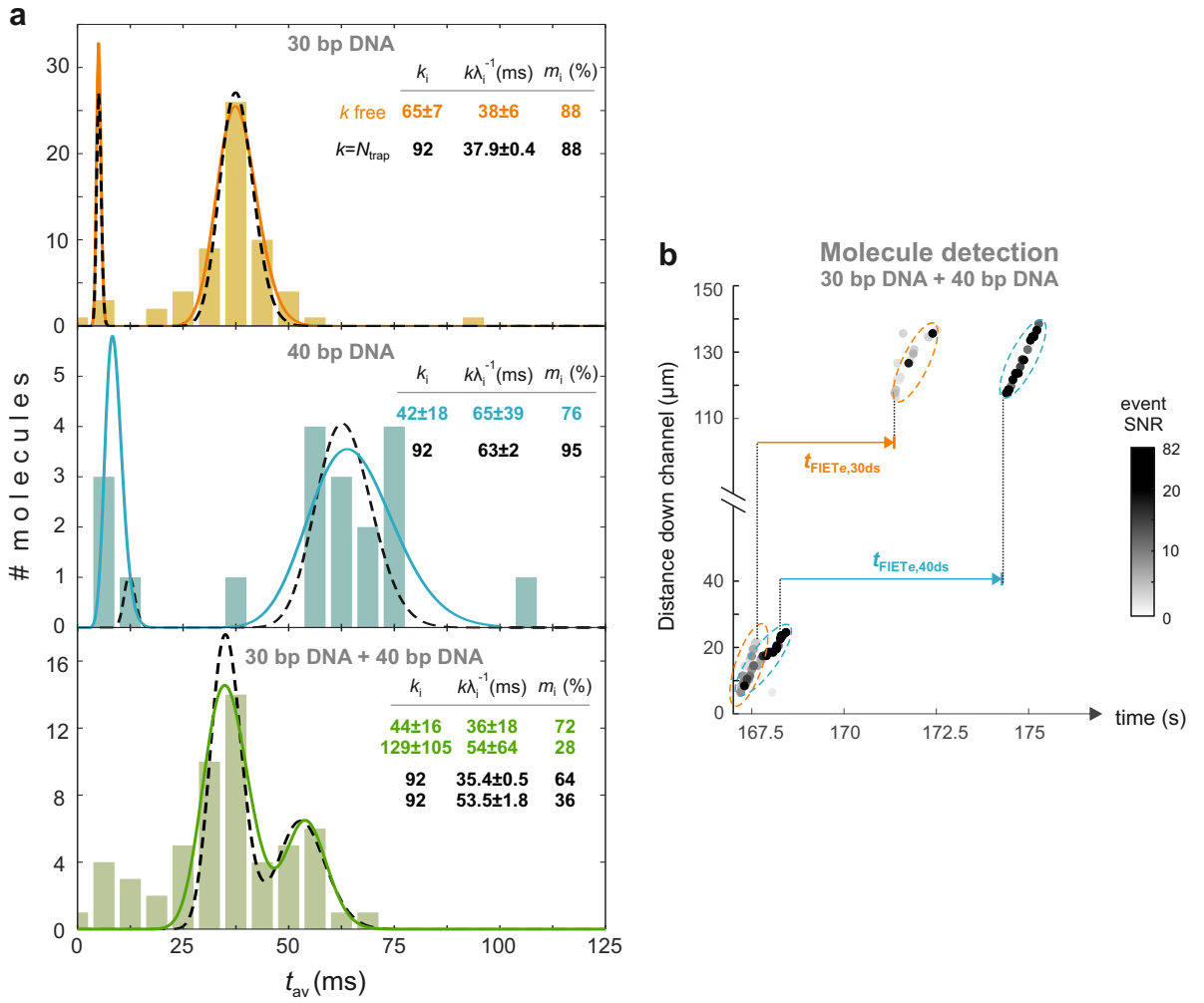


Fig. 16 **Measurement of 30 and 40 bp DNA using FlET*e*.** (**a**) t_{av} distributions from measurements of 30 bp, 40 bp and a 1:1 mixture of the two DNAs measured in a device with $2h$=70 nm. A bin width corresponding to average uncertainty on t_{av} for the 40 bp DNA spectra was used for constructing all histograms, 6.25 ms. The distributions are fit with a sum of Erlang distributions (Eq. 16). For the pure samples the fit parameters for the majority fraction are shown (coloured lines) and we see broadening of the distribution compared to the theoretical limit (black dashed lines). There was a consistent fraction measured with low t_{av} in all measurements which could be indicative of sample heterogeneity such as free dye. For the mixture we see two peaks with mean values, $k\lambda^{-1}$, which match the pure samples. The amplitudes of the two distributions depart slightly from the 50% expected from the input concentrations. (**b**) Exemplary single molecule traces showing the detection of two molecules from the mixed sample, showcase the difference in $t_{\mathrm{FlET}e}$ for a 30 bp vs 40 bp DNA who were in the start observation window coincidentally.

Methods to improve the throughput of FlET*e* experiments

The current throughput is limited by the low sample concentration required for molecule identification. Introduction of a short "plug" of many molecules would mean molecules only need detecting once at the finish observation window since their start time is known. This innovation would also relieve the pressure on N_{trap} since no observation window is required at the start of

the channel, so an arbitrary length of channel can be used for the measurements and only a small FOV microscope would be required.

The length of measurements currently is also an obstacle to throughput, with manual input required for 10s of minutes of recording to keep the device in focus against stage drift. The implementation of an autofocus system would allow for automated measurements over a longer time period to measure more molecules.

Methods to improve the resolution of FlET*e* experiments

As mentioned previously, precision and hence resolution of FlET*e* measurements is limited by N_{hop}. Pursuance of the alternative microscope design with two decoupled illuminated and imaged sections of the device arbitrarily separated, would allow a large increase in N_{trap}. Alternatively, a new device design involving a single curved channel per field of view (e.g. snaking from an observation window in the top left corner to another in the bottom right corner) would provide a far increased N_{trap}. Either of these options combined with the addition of positive pressure to the upstream side of the device as well as negative, would allow for increased flow speeds and hence L can be reduced for a given Pe. Alternatively electro-osmotic flow could be utilised relieving the linear decay of velocity with channel length as we have for pressure-driven flow[100], although this brings with it the addition of electrophoretic molecular motion, which could interfere with the electrostatic trapping.

Chapter 4

Escape time size-and-shape spectrometry (ETsy)

There is considerable interest in probing biomolecular properties under physiological conditions. This chapter addresses the considerations required of ET*e* technology when applied to the high salt and neutral pH inherent of physiological conditions, and measurements of molecular size and shape are showcased.

4.1 Escape-time measurements under physiological conditions

In order to use ET*e* technology in physiological conditions several key steps were required to be made. Firstly, sensitivity to molecular properties is required under these conditions. From Eq. 10 it is apparent that high salt concentration greatly reduces the magnitude of the electrostatic interaction energy with the free-field fluid trap. However, as noted in section 1.4.3, the translational contribution ΔF_{trans} to the fluctuation free energy f, still offers significant well depth, due to the contribution of entropy, see Eq. 20.

$$\Delta F_{\text{trans}} = k_{\text{B}} T \ln\left(\frac{h_2}{h_1}\right) \tag{20}$$

Where we now redefine the channel height as h_1 (=$2h$ previously), and the depth in the trap to be h_2 (=$2h$+d previously). As the channel height approaches the dimension D, of a molecule, we are required to consider higher order terms in the expression, which come from the excluded

volume reducing the number of accessible states available to the molecule, Eq. 21.

$$\Delta F_{\text{trans}} = k_{\text{B}} T \ln \left(\frac{h_2 - D}{h_1 - D} \right) \tag{21}$$

When combined with Kramers' exponential sensitivity of escape time on an energy well depth (Eq. 11) we see a non-linear dependence of t_{esc} on the size of the molecule, Eq. 22. We use this as our physical read-out of the molecule under physiological conditions.

$$t_{\text{esc}} = t_{\text{r}} \left(\frac{h_2 - D}{h_1 - D} \right) + \text{c} \tag{22}$$

In order to achieve the shall channel heights of the order of molecular size, an improved fabrication procedure was developed using thermal bonding to avoid the sagging of glass lids to the devices upon anodic bonding as previously reported[93]. More details is outlined in Appendix A.

We note that the sensitivity of t_{esc} to D, although non-linear, is not as high as the exponential dependence on q_{eff} in the lower salt ETe measurements. To overcome this, higher event numbers are required to reduce the statistical uncertainty on t_{esc}. This was achieved using an upgraded microscope with 20-fold increase in the FOV, thus highly parallelising the measurement by a similar factor. 100k events are now achievable in a <10 min measurement, yielding a <1% imprecision on t_{esc} with the new microscope. Appendix B details the improvements.

Another "sticking point" with measurements in physiological conditions is the increase in non-specific binding of molecules to the channel surfaces, so-called sticking. This is more likely under the low electrostatic environments of higher salt concentrations and lower pH, as there is less electrostatic repulsion away from the channel walls. Sticking is a phenomena widely reported in the literature[101], which comes about from Van der Waals forces of attraction from interactions between the electron clouds of the molecule and the channel wall[102]. To overcome this, many passivation methods have been reported including lipid bilayers, poly-electrolytes and polymer coating of the surface. We adopted the latter due to the ease of silanisation of the silica surfaces of the device with a methoxy polyethylene glycol silane (mPEG-silane).

These components combine to enable physiological ETe measurements of single-molecule size and shape which we term escape-time size-and-shape spectrometry, ETsy. We exemplify ETsy with a series of structural and kinetics measurements of nucleic acids and folded and disordered

proteins, described in the following paper: "Single-molecule size-and-shape spectrometry on a chip".

4.1.1 Single-molecule size-and-shape spectrometry on a chip

Briefly, we show ETsy measurements of the molecular size of proteins over 3 orders of magnitude of molecular weight, and with the precision to distinguish small molecules just two carbon atoms different in mass. We showcase the shape-sensitivity of ETsy with measurements of the maximum dimension of two DNA nanostructures, and of the rise per base pair of DNA and RNA helices. When combined with the single-molecule tracking employed in the TrET*e* measurements in section 3.1, we measure the complex series of conformations present in a riboswitch sample, and shed light on the interconversion rate of the sub-populations. We use the short measurement times to probe chemical and biomolecular interaction kinetics and binding affinties, vetted with a series of DNA hybridisation reactions over six orders of magnitude in affinity. We apply this to the detection of insulin at near physiologically relevant concentrations, based on the conformation change induced in the insulin receptor upon binding.

We envisage a host of further studies employing ETsy for both analytical and diagnostic purposes, building upon these measurements outlined in the manuscript.

The supplementary materials for this paper can be found in Appendix H.

Title: Single-molecule size-and-shape spectrometry on a chip

Authors: Xin Zhu[1]†, Timothy J. D. Bennett[1]†, Konstantin C. Zouboulis[1,2], Dimitrios Soulias[1], Michal Grzybek[3], Justin L. P. Benesch[1,2], Afaf H. El-Sagheer[4,5], Ünal Coskun[3,6-8], Madhavi Krishnan[1,2]*

Affiliations:

[1]Physical and Theoretical Chemistry Laboratory, Department of Chemistry, University of Oxford; South Parks Road, Oxford OX1 3QZ, United Kingdom

[2]The Kavli Institute for Nanoscience Discovery, University of Oxford; Sherrington Road, Oxford OX1 3QU, United Kingdom

[3]Center of Membrane Biochemistry and Lipid Research, University Hospital and Faculty of Medicine Carl Gustav Carus, Technical University Dresden; 01307 Dresden, Germany

[4]School of Chemistry and Chemical Engineering, University of Southampton; Highfield, Southampton, SO17 1BJ, United Kingdom

[5]Institute for Life Sciences, University of Southampton; Highfield Campus, Southampton, SO17 1BJ, United Kingdom

[6]Paul Langerhans Institute Dresden (PLID) of the Helmholtz Center Munich, University Hospital and Faculty of Medicine Carl Gustav Carus, Technical University Dresden; 01307 Dresden, Germany

[7]German Center for Diabetes Research (DZD), 85764 Neuherberg, Germany

[8]Max Planck Institute of Molecular Cell Biology and Genetics, 01307 Dresden, Germany

*Corresponding author. Email: madhavi.krishnan@chem.ox.ac.uk

† These authors contributed equally to this work

Abstract: Size and shape are critical discriminators between molecular species and states in solution. We have developed a micro-chip based high-throughput imaging approach offering rapid and precise determination of molecular properties under native conditions. Our method detects differences across at least three orders of magnitude in molecular weight and down to a carbon atom in small molecules. We quantify the strength of intermolecular interactions over 6 orders of magnitude in affinity constant, and track reactions in real-time. Single-molecule spectra of biomolecular conformation serve to characterize sample state heterogeneity at the highest resolution, offering predictive input to model 3D structure. We further leverage the power of this structural sensitivity for diagnostics, exploiting ligand-induced conformational changes in the insulin receptor to measure insulin concentration in serum-like conditions at the sub-nanoliter and sub-zeptomole scale.

One-Sentence Summary: A micro-chip discriminates between molecular species and states based on shape and size, offering a new analytical and diagnostic technology.

Main Text:

Experimental techniques that deliver high resolution characterization of biomolecular properties in the solution phase are of great current interest. A range of highly sophisticated and powerful approaches focus on the measurement of molecular properties such as size, mass, charge, dipole, and shape or 3D conformation, with the underpinning mechanisms generally involving externally applied fields, e.g., electrical (*1, 2*), and magnetic fields (*3, 4*), X-rays (*5, 6*) and light (*7-14*), as well as the use of external fields in conjunction with sieving matrices and micro- or nanoscale transport principles (*15-26*). However, surface interactions, the presence of matrices, as well as strong electrical fields and radiation can perturb the integrity and function of biomolecules whose 3D conformation generally arises from the concerted action of weak intramolecular forces (*27, 28*). Therefore, characterization methods that offer high-precision readouts whilst preserving molecular structure and properties are of intrinsic value.

The friction associated with an object moving through a fluid carries a clear signature of its volume and shape. In random thermal motion, the larger the object is, the shorter the average distance it travels in a given period of time. The Stokes' or hydrodynamic radius, r_{H}, quantifies this property, and has been extensively used to characterize molecular species in solution (*29-32*). The shape of an object may be further captured in terms of its compactness given a fixed volume, as reflected in the diameter of the smallest bounding sphere, D_{s}, enveloping it (Fig. 1). Indeed, workhorse laboratory techniques such as size-exclusion chromatography exploit this molecular property in order to separate and sieve mixtures of species. Here, the spatial extent of the molecule determines the length of the path it takes, and therefore the speed at which it traverses a porous landscape. The diameter of the 'molecular envelope' D_{s} is further reminiscent of D_{max}, the maximum spatial extent of the molecule, encountered in small-angle x-ray scattering (SAXS) (Fig. 1C). Precise measurement of the size and shape of molecules achieved by direct observation of molecular motion in a suitably tailored landscape could therefore hold promise for highly simplified structure-sensitive readouts traditionally believed to lie exclusively within the realm of high-resolution techniques (Fig. 1A) (*33-35*).

In this work we leverage the combined advantages of random, thermally driven motion and size-dependent confinement effects at the nanoscale to achieve high measurement sensitivity to molecular conformation under native conditions in the solution phase. Spatially constraining the vertical motion of an object diffusing in a geometrically tailored slit can dramatically enhance the time the object spends in local, deeper, 'pocket' or trap regions, in a size-dependent fashion (Fig. 1A, B, Movie S1). The larger the object, or the greater the degree of confinement in the parallel-plate slit region, the longer it resides in the trap. Engineering the number of states accessible to a molecule in the slit – e.g., translational, configurational and conformational – entropy in its various forms can be exploited at the nanoscale to significantly amplify the impact of size and shape on molecular exit (escape) times. Low-power imaging of the motion of thousands of molecules in a nanostructured landscape, using a simple wide-field optical microscope, thus delivers a rapid and precise multi-parametric readout of molecular size and shape (Movie S1). We term the approach escape time size and shape spectrometry (ETsy) and demonstrate the broad ramifications of the method's core working element – the 'entropic fluidic trap'– in a range of analytical application areas, including molecular diagnostics (*21*).

Spatial tailoring of translational entropy enables high throughput measurements and molecular counting

In the experiments reported here we introduce fluorescent molecules at a concentration of 10 fM-10 nM suspended in phosphate buffered saline (PBS, ionic strength ≈160 mM) into a series of parallel-plate slits of height h_1 that carry arrays of periodic cylindrical indentations of total height h_2 (Fig. 1A, B). Similar to previous work, the "pocket" nanostructures act as thermodynamic traps for single molecules in solution (Fig. 1B) (*20, 34-36*). In contrast to previous work, charge-related electrostatic effects are absent, and the trap is purely entropic (*21*). We image the diffusive dynamics of molecules, labelled on average with a single fluorophore, using standard wide-field microscopy as described previously (Fig. 1A). We determined the average escape time, t_{esc}, for a variety of molecules from small organic fluorophores to disordered and globular proteins, nucleic acids and protein complexes by imaging the chip for about one minute in each case (Fig. 1D, E and 2A). Typical imaging conditions entail exposure times $t_{\mathrm{exp}} = 5$ ms and an imaging rate of 100 Hz, unless stated otherwise (see supplementary text for details).

For a small sphere whose radius $R \approx r_{\mathrm{H}} << h_1$, we have $t_{\mathrm{esc}} \propto r_{\mathrm{H}}(h_2/h_1)$, which illustrates that the height modulation in the slit enhances molecular residence times by a factor h_2/h_1 compared to the free-diffusion value (red line, Fig. 2C, fig. S7). This entropic enhancement of residence times in the pocket region is crucial to the method since it fosters high throughput direct imaging of the escape process on the order of 100-1000 molecules in a minute or less of measurement, using both low incident optical power densities (~5 W/mm^2) and a large-area detector such as a camera chip. Our statistically dominated measurement imprecision depends on N, the number of escape events observed, as $1/\sqrt{N}$. Large values of $N = 10^4 - 10^5$ achieved using highly parallel observation of the escape process at a molecular concentration of about 100 pM, implies $\leq 1\%$ measurement imprecision on r_{H}, and is key to high precision measurement. High throughput and speed are also critical to achieving a low limit of quantification in a method that relies on counting a specific molecular species rendered visible by labelling, against a complex yet optically silent background. The present experimental configuration offers quantitative detection of a labelled species down to ~10 fM in a 10 min continuous-flow measurement, with opportunity for increasing throughput and lowering the detection threshold by at least a factor 100 through further optimization of device design (fig. S5).

Measuring t_{esc} on a range of molecules spanning 3 orders in molecular weight, we noted the scaling $t_{\mathrm{esc}} \propto r_{\mathrm{H}} \propto M^{1/3}$ as expected for globular molecules, which implies that $< 1\%$ imprecision on t_{esc} translates to $< 3\%$ measurement imprecision in molecular weight (Fig. 2A). By contrast, for disordered proteins, e.g., Prothymosin-α (ProTα) and Starmaker-like protein (Stm-l) we measure larger escape times that are in agreement with the scaling relationship $t_{\mathrm{esc}} \propto r_{\mathrm{H}} \propto M^{3/5}$ characteristic of unstructured polymers (Fig. 2A, fig. S8) (*4, 6, 37*).

For small molecules of molecular weight, $M \lesssim 1$ kDa, $\approx 0.5\%$ measurement imprecision on the escape time, t_{esc}, implies the ability to detect differences of at least a single carbon atom between species as shown in Fig. 2B (fig. S8). We illustrate the power of both the speed and precision of the approach in real-time tracking of the change in molecular weight of an ester due to hydrolysis in aqueous solution (Fig. 2B, lower panel). The N-hydroxysuccinimide (NHS) ester of the fluorescent dye ATTO532 hydrolyses in water to form carboxyl-ATTO532 and the leaving NHS group, with the rate of hydrolysis strongly accelerated under alkaline conditions (*38*). We performed continuous-flow monitoring of the escape process for an aqueous solution of ATTO532-NHS ($M = 743$ Da, fig. S1) at pH 7 and pH 9. We determined the average escape time

t_{av} over all recorded escape events of duration Δt using a series of intermittent measurements of 30 sec duration performed over a total experimental period of about 10 minutes. A systematic overall reduction of 3% in measured t_{av} over the course of the reaction implies a mass change of about 100 Da, commensurate with that of the NHS leaving group. The measured decay rate constant of 2.3×10^{-3} s^{-1} at pH 9 agrees well with literature values (*38*).

Nanoscale confinement offers non-linear sensitivity of the readout to molecular size and shape

Importantly we may use the measured t_{esc} values to quantitatively infer molecular hydrodynamic radii, r_{H}, and bounding sphere diameters, D_{s}. In order to do so we performed Brownian Dynamics (BD) simulations of the escape process for spheres of radius r_{H} and $D_{\mathrm{s}} = 2r_{\mathrm{H}}$ as described previously (*39*). Simulated average escape times indeed display good agreement with the simple model outlined in the supplementary text and may be encapsulated in the following equation (*39*) (fig. S7):

$$t_{\mathrm{esc}} = A r_{\mathrm{H}}(h_2 - D_{\mathrm{s}})/(h_1 - D_{\mathrm{s}}) + t_{\mathrm{o}} \qquad (1)$$

We find that a coefficient of $A \approx 0.3$ s/μm that accounts for the viscosity, device geometry and various aspects of the imaging-based detection process, as well as the inclusion of an offset $t_{\mathrm{o}} = 5.83$ ms $\approx t_{\mathrm{exp}}$ quantitatively fits the simulated escape times. Thus, in a typical experiment to determine r_{H} values, we first use globular proteins of known r_{H} to determine h_1 from fits of the data to Eq. (1). We then use this calibration function to convert the measured t_{esc} for a test molecular species to hydrodynamic radius r_{H} (fig. S9). Indeed r_{H} values for a set of globular proteins measured in slits of calibrated height $h_1 = 23.8$ nm agree well with values computed based on molecular structures using the program HYDROPRO (*30*) (Fig. 2D, table S4).

Importantly, Eq. (1) suggests that, owing to rotational diffusion, molecules of non-spherical shape may exhibit a further enhancement of t_{esc} compared to their spherical counterparts of similar hydrodynamic radius. Since rotational times are generally much faster than translational diffusion, we may assume that a molecule that rotates isotropically while translating sweeps out a local volume defined by a sphere of minimum diameter $D_{\mathrm{s}} > 2r_{\mathrm{H}}$ as shown in Fig. 1C. The implications of this high sensitivity to the molecular envelope are illustrated in our ability to measure differences of ≈0.3 bp (≈2 Å) between rod-shaped double-stranded (ds) DNA fragments in the size range n_{bp} =56-60 bp (≈ 20 nm in length) (Fig. 2E).

To test this further we performed measurements on dsDNA and dsRNA of length n_{bp} =30 to 60 basepairs, entailing lengths in the range 7 to 21 nm. Performing measurements of t_{esc} in slits of depth $h_1 \approx 25$ nm we found a progressive non-linear increase in t_{esc} with r_{H} values of the molecular species compared to the expectations for spheres of the same hydrodynamic radius r_{H} (fig. S10). Fitting the measured t_{esc} vs. n_{bp} data using Eq. (1), assuming $D_{\mathrm{s}} = b n_{\mathrm{bp}}$, and using theoretical expressions for r_{H} for cylinders we obtained a rise per basepair $b \approx 3.2$ Å for the B-form double helix characteristic of dsDNA and $b \approx 2.3$ Å for the A-form double helix expected for dsRNA (Fig. 3A, fig. S10). These values are remarkably close to those inferred from high resolution structural measurements and previous solution phase structural studies (*34, 40, 41*). The high sensitivity of the technique to the radius of the bounding sphere that serves as a proxy for molecular shape may be further harnessed towards detecting minor alterations in conformation resulting from folding or binding processes, or to detect the presence of rare states or complexes that entail changes in size and/or shape, as described later.

We note that in this study we focus on analysing the response expected for stably folded molecules in the regime $D_s < h_1$. However, for molecules described by $D_s > h_1$, e.g., disordered molecules whose maximal spatial extent exceeds h_1, we may expect an even greater enhancement of escape times arising from the contribution of conformational entropy to the readout. This mode of measurement representing 'extreme confinement' may be investigated in future work.

Inferences on DNA-nanostructure size and shape

Since we expect ETsy to offer strong and tunable shape-based discrimination of molecules that may be effectively identical in all other relevant respects (e.g., in hydrodynamic radius, mass and charge), we studied DNA nanostructures that carry almost exactly the same number of basepairs but differ in 3D conformation (*36*, *42*, *43*). We considered DNA nanostructures composed of 240 bases in states referred to as 'bundle' and 'square-tile' as previously described (*36*) (Fig. 3B), where $D_s \approx 20$ nm for the bundle provides an estimate of the size regime of the measurement. The 'bundle' and 'tile' are experimentally indistinguishable from each other in diffusion coefficient measurements by FCS (table S5) (*36*). We performed t_{esc} measurements on both species in slits of height $h_1 \approx 100$, 70 and 40 nm. Fig. 3B demonstrates the ability to use increasing confinement (lower h_1) in order to accentuate disparities in t_{esc} that arise on account of 3D conformational differences, reflected in D_s, between the two nanostructures, as suggested by Eq. (1). The relative insensitivity of t_{esc} to molecular conformation in the regime $h_1 \geq 70$ nm largely reflects a response expected based on pure diffusion alone. However, we note systematically increasing t_{esc} values, with disparities between the two DNA nanostructures reaching a factor 1.2 as h_1 approaches D_s. This underscores the fact that the measurement mode may be tuned to enhance the sensitivity of the readout to molecular 3D conformation.

Furthermore, according to Eq. 1, measurements performed using two different values of h_1 ought to support a readout of two properties of the molecule, namely r_H and D_s. We used t_{esc} measurements in slits of two different mean calibrated height values of h_1 =41.4 nm and 73.9 nm in order to infer these conformational parameters for the 'bundle' and 'tile' (fig. S11). Fig. 3C provides a graphic representation of the procedure behind such an inference. A pair of independent measurements using sufficiently different h_1 values yields two curves on a plot of D_s vs. r_H whose point of intersection reflects the r_H and D_s values describing the molecular state. Comparing our inferences with values expected from molecular structural models obtained from coarse grained MD simulations using oxDNA, and verified using SAXS (*43*, *44*), indeed reveals good mutual agreement (table in Fig. 3C, fig. S11; see supplementary text for detail). For example, whilst the measured r_H values for the 'tile' and 'bundle' are rather similar, their D_s values are clearly different and in fact appear to be entirely responsible for the measured disparity between the two species. Next, we demonstrate the use of this sensitivity to molecular shape to examine conformational heterogeneity in folded nucleic acid species such as RNA.

Resolving mixtures of molecular conformational states by constructing single-molecule spectra

Escape time data examined at ensemble level as shown in Figs. 1-3 can sometimes reveal the presence of multiple escape timescales, depending on the sample, (fig. S8; see supplementary text for detail). A multitude of timescales generally contains important information on the sample, e.g., heterogeneous states of multimerization, or distinguishable 3D conformation states in monomeric molecules. Extracting these time components and their abundances in ensemble measurements can under certain conditions be subject to mathematical difficulties (*45*). Accessing information at the level of individual molecules not only avoids potential ambiguities of multi-exponential fitting,

but can also furnish insight into the nature of the underlying states, e.g., whether they stem from individual molecules characterised by distinct 3D conformations that are stable over the observation window, or rather from faster interconversion between different conformational states. Because we observe the hopping dynamics of every individual molecule in the microscope's field of view, we may indeed construct spectra of single-molecule properties in a mixture of different molecular species or states as shown in Fig. 4. Performing measurements at low sample concentrations ($\sim$10 pM) permits facile recognition of trajectories arising from individual molecules whose average escape time now describe the properties of individual molecular states (Fig. 4A, Movie S2, see supplementary text for detail).

We first validated the ability to identify, distinguish and quantify stable single-molecule states by examining an equimolar mixture of 24 bp and 60 bp DNA: two geometrically highly well-defined and conformationally stable species (Fig. 4B, fig. S14). We constructed single molecule spectra by recording the average escape time t_{av} for a number of single-molecule hopping trajectories each consisting of N_{hop} escape events of duration Δt. Each measurement in the resulting molecular escape spectrum thus represents an average value originating from a separate molecule. We currently record about $N_{\mathrm{hop}} = 200 - 300$ hops on average from a molecule before either the label photobleaches or the molecule diffuses out of the field of view. Such a measurement currently gives a typical $\sim$3-5% measurement imprecision on t_{av} per molecule, which can be significantly improved upon in future. The mixture of dsDNA species, indeed displayed two peaks in the molecular escape-time spectrum at mean molecular t_{av} values given by μ_1 and μ_2, that were in agreement with the two escape timescales t_1 and t_2 obtained from a bi-exponential fit of all the data pooled together in the ensemble approach (bottom row, Fig. 4B). Furthermore, the fractional abundances of the two species, $m_1 \approx m_2 \approx 50\%$, were also in good agreement with the ensemble fits.

3D conformation of a riboswitch in aqueous solution

We then applied this single-molecule trajectory analysis to assess the 3D conformation of folded RNA. Like proteins, RNA can fold into well-defined 3D structures to carry out a wide range of cellular functions, but structure determination and prediction for RNA has significantly lagged behind their protein counterparts (*46*). Riboswitches are an important class of folded RNA molecules that sense small molecules and are involved in the regulation of gene expression. In bacteria, e.g., riboswitches can regulate sulfur metabolism by binding to the small molecule S-adenosylmethionine (SAM), offering potential targets for new antibiotics. In fact the structure of a $\sim$40 kDa SAM-IV riboswitch (fig. S4), one of the three most commonly found riboswitches, has only recently been solved by cryogenic electron microscopy (cryo-EM) (pdb codes: 6UES and 6UET) (*47*). We therefore examined the possibility of using ETsy to detect properties of the solution phase 3D conformation of the SAM-IV riboswitch singly labelled with ATTO 532 (fig. S12, S16).

Interestingly, rather than a single timescale corresponding to the predominant cryo-EM structure (70% abundance, see Fig. 4C), we found that both ensemble-averaged measurements and single molecule spectra of the SAM-IV riboswitch, revealed three distinct escape timescales (Fig. 4D-F, fig. S16). Examination of individual molecular trajectories revealed average molecular escape times falling into three distinct classes over the ca. 6 sec measurement window, indicative of stable molecular conformations rather than rapid interconversion between states (Fig. 4 D-F, fig. S15). In order to benchmark the observed timescales, we performed control measurements on free fluorescent dye, 30 bp dsDNA and a 119 nt single-stranded (ss) DNA all of which entail little

or no secondary structure (Fig. 4C, D, lower panels; fig. S16). The smallest timescale, $\mu_1 \approx t_1 = 12$ ms, in both the single molecule spectrum and the ensemble measurements of SAM-IV sample can be attributed to free fluorescent dye molecules in the mixture which are generally difficult to remove entirely. However, we attribute the second and third measured timescales to two different conformational states, or similar groups of states, of the SAM-IV molecule – one compact and the other more extended (*5*). The larger of the molecular timescales, μ_3, is similar to the timescale measured for 30 ds DNA whose value of $D_s \approx 9.6$ nm in fact closely resembles that of the reported 6UES structure (*47*) (Fig. 4D). The smaller of the two observed SAM-IV timescales, μ_2, in turn is indicative of a conformation of ~60% abundance, suggesting an overall more compact and globular state. All mean timescales at the ensemble- and single molecule-level are in good mutual agreement, i.e., $t_i \approx \mu_i$ (Fig. 4G). Fig. 4G displays a summary of the conformation analysis where the main conclusions are that (1) the measurements $t_3 \approx \mu_3$ (blue symbol and shaded zone in Fig. 4G) are consistent with that of an oblate ellipsoid characterized by D_s and r_H values that are very similar to those computed for the 6UES cryo-EM structure (black cross in Fig. 4G), and (2) the observation $t_2 \approx \mu_2$ indicates a conformation characterized by an ellipsoid of smaller D_s and r_H compared to 6UES (red symbols and shaded region)

Furthermore, upon addition of a high concentration of the SAM ligand, the escape time spectra remained comparatively unchanged suggesting no significant conformational change of the riboswitch upon ligand binding (fig. S16). Although surprising at first glance, this observation is in fact in line with cryo-EM data as well as with observations on other riboswitch aptamers, e.g., SAM-I and lysine riboswitches that do not change conformation upon ligand-binding (*48, 49*).

Measuring binding affinities and rates of intermolecular association and dissociation

Nucleic acid interactions

Having explored the ability of the technique to read out molecular weight and conformational states of purified molecular species, we next explore the capability of our approach in making rapid, quantitative measurements of intermolecular interactions in mixtures, as reflected in measured dissociation constants, K_d, as well as of on- and off-rates of binding – k_{on} and k_{off} (Fig. 5). DNA hybridization provides a simple tuneable system where intermolecular binding affinities can be varied substantially using single basepair increments between complementary stretches, supporting comparisons of measurements not only with independent experimental techniques but also with theoretical estimates. We validated the measurement approach using fluorescent molecular 'bait' represented by short strands of ATTO532-labelled ssDNA ranging in length from 7-15 bases hybridizing to the corresponding complementary sequence at the 5' end of an approximately 200 base long "scaffold" of unlabelled ssDNA (Fig. 5A, B). Mixtures of the labelled oligo at a concentration of 0.02-1 nM and unlabelled DNA at various concentrations are incubated to equilibrium, loaded into the trap landscape and escape times recorded over a period of 1-5 min. Histograms of event durations reveal two dominant timescales: a short t_1 ≈10 ms corresponding to the free labelled oligo and a longer t_2 ≈30 ms corresponding to the oligo bound to the 200-base ssDNA (Fig. 5A). The fractional abundances of the two species $m_{1,2}$ are determined using bi-exponential fits and plotted as a function of the concentration of the unlabelled binding partner to determine the K_d value of the interaction (upper panel, Fig 5B, fig. S17; see supplementary text).

We observed a systematic increase in binding affinity with increasing oligo length, reflected in a progressively decreasing measured K_d values, in good agreement with previous reports (Fig. 5D) (*50*). Note that although samples are incubated over timescales ranging from several hours to several days in order to ensure that the binding reaction reaches equilibrium, the $t_{1,2}$ measurements

corresponding to one point on a plot displayed in Fig. 5B takes about 1-5 min over the range of concentrations of unlabelled binding partner probed (fig. S17). Furthermore, the speed of the readout can be harnessed to perform real-time monitoring of the sample following either the mixing of the two unbound species, or alternatively, following the addition of an unlabelled "chaser" binding partner to the bound complex (lower panel Fig. 5C; fig. S18) (*51*). Measuring the fractional abundances as a function of time, similar to Fig. 2B, permits the measurement of off-rates and corresponding on-rates for the same reaction. We find that these measured rates yield measures of $K_{\mathrm{d}} = k_{\mathrm{off}}/k_{\mathrm{on}}$ that are consistent with the equilibration-based measurement approach as well as with theoretically expected values (*50, 52*) (Fig. 5C).

DNA-protein and protein-protein interactions

We further examined DNA-protein binding affinity measurements performed with a singly labelled dsDNA molecule binding to unlabelled restriction enzymes (Fig. 6A, fig. S20) (*53*). Here, we added the unlabelled binding partner, the restriction enzyme EcoRI, at various concentrations to a fixed concentration of 0.25 nM of fluorescently labelled 24 bp DNA containing the enzyme recognition site. In general, binding systems that produce large shifts in t_{esc} of a factor 2 to 5 upon binding yield $P(\Delta t)$ histograms that are amenable to multi-exponential fitting and the extraction of fractional abundances of the underlying species as (Fig. 5A). We determined the bound fraction of DNA from fitting two timescales to the t_{esc} data as described in Fig. 5, and obtained a value of $K_{\mathrm{d}} \approx 1.5$ nM. However, many binding reactions are associated with small changes in size and or overall conformation of the complex, which implies small changes in t_{esc}, sometimes as small as 10% or less. Here, the similarity of timescales corresponding to the two molecular states can render robust extraction of two exponential components and their individual fractions from the escape time data untenable (*45*). Therefore, rather than fitting escape time histograms we track the quantity t_{av}, which represents a simple average over all escape events of duration, Δt, in the mixture of the two interacting species. We applied this procedure to examine a well-characterized antigen-antibody interaction between human leukocyte antigen (HLA) isoform A*03:01 (HLA A*03:01) and W6/32, a mouse-derived pan-HLA class 1-reactive monoclonal antibody (fig. S3) (*54*). Anti-HLA antibodies are important as clinical biomarkers for evaluation and monitoring of allograft rejection risks in cell and organ transplantation (*55*). Binding of the labelled HLA antigen ($M \approx 49$ kDa) to its antibody ($M \approx 150$ kDa) results in a 40% increase in t_{av} relative to that of the free, labelled bait. Analysing the t_{av} data to determine the bound fraction, we obtained $K_{\mathrm{d}} \approx 0.4$ nM for this interaction, in good agreement with reported values of $K_{\mathrm{d}} \approx 0.7$ nM (Fig. 6B, fig. S20) (*56*). Furthermore, our measurement displays sensitivity to the HLA Antibody at concentrations as low as < 100 pM, providing a general approach to detecting antibodies in solution. The measurements also shows that t_{av} alone can serve as an excellent quantitative measure of the fraction of fluorescent bait in the bound-state, and this approach is utilized again in investigations involving direct detection of free insulin in solution, described next (fig. S19; see supplementary text for detail).

In order to explore the possibility of detecting a small, medically relevant protein analyte such as free insulin in a complex sample, we first examined the use of a fluorescently labelled 31-base insulin-binding aptamer that is expected to bind insulin in the folded G-quadruplex containing state (*57*) (Fig. 6C). Upon formation of the structured state, ETsy directly detects a 10% reduction in t_{av} for the aptamer (Fig. 6D, fig. S21). We incubated 5 nM of singly-labelled, folded aptamer with 100 μM unlabelled insulin which is well above the reported K_{d} of the interaction (*58, 59*). We noted a 15% increase in t_{av} for the folded aptamer which may be expected based on the increase in mass of the complex alone. Importantly we noted no change in t_{av} for the unfolded

aptamer in the presence of insulin, pointing to the apparent requirement of secondary structure in the aptamer for insulin binding (*57*). Measuring t_{av} for mixtures containing variable concentrations of free, unlabelled insulin and a fixed low concentration of labelled aptamer 'bait' of 1 nM, we obtained a $K_{\mathrm{d}} \approx 90$ nM, in broad agreement with literature values (*58, 59*). This measurement demonstrates a straightforward, rapid and general approach to detecting the presence of a small, medically important analyte in solution. However, the relatively modest affinity of the aptamer for insulin does not hold significant promise for direct insulin sensing at physiologically and biochemically relevant concentrations.

We therefore set out to characterize a higher affinity interaction for insulin using a biologically relevant binding partner that is tuned to respond to low nM physiological concentrations of insulin. The insulin receptor (IR) is a transmembrane receptor tyrosine kinase that is activated by insulin binding to its extracellular ectodomain (ECD) (*60*). We measured the binding affinity between a singly labelled insulin molecule ($M \approx 6$ kDa, fig. S2) and the soluble IR-ECD ($M \approx 280$ kDa) by ETsy. The large disparity in timescales between free and bound insulin facilitates a direct measurement of bound fractions using two fixed timescales in the ensemble fit (Fig. 6C, fig. S20). The highest affinity we measured for the interaction was $K_{\mathrm{d}} \approx 7$ nM which is smaller than the previously reported K_{d} ≈30 nM value measured for the same IR-ECD by nano–differential scanning fluorimetry (*60, 61*), and closer to the sub-nM high affinity binding of the first ligand binding site of the full-length insulin receptor (*62-64*).

Structural inferences on the 'Apo' and 'Liganded' states of the insulin receptor ectodomain

Having established the ETsy method for affinity measurements on IR-ECD and labelled insulin we explored the general potential our platform for membrane protein receptor-based immunoanalysis and diagnostic applications. We therefore turned our attention to using fluorescently labelled IR-ECD as 'bait' to detect the much smaller, unlabelled insulin molecule in solution. IR-ECD is expected to bind up to 4 insulin molecules, entailing an increase in molecular weight of ca. 9% ($\Delta M \approx 25$ kDa) and a modest concomitant increase of ≈3% in t_{esc} (or r_{H}) (*61*).

In order to examine the conformational properties of the apo- and liganded IR-ECD we first carried out a structure/conformation analysis of IR-ECD in the insulin-free and bound states using ETsy as described in Figs. 3 and 4. Based on structural data from cryo-EM, the soluble ectodomain (ECD) of the insulin receptor is expected to have an approximate spatial extent of 17 nm, a hydrodynamic radius of $r_{\mathrm{H}} = 6$ nm, and can bind up to 4 insulin molecules. Molecular structures obtained using X-ray crystallography and cryo-EM indicate that upon insulin binding, the IR-ECD undergoes a significant conformational change from an 'inverted-V' apo-state to a fully liganded 'T-state' including intermediate states with a lower number of bound insulin molecules (*61, 65*), as also reported for snail venom insulin:human IR-ECD interactions (*66*) (pdb codes: 4ZXB 'apo', 7QID 'intermediate', 6SOF 'fully liganded') (Fig. 7A). However, the generation of crystal structures of IR-ECD in the 'apo' state has relied on internal conformational stabilization. Indeed the high degree of conformational flexibility of the apo-state compared to its liganded counterpart has precluded the identification of a representative 3D structure for the ligand-free state using cryo-EM (*61*).

We performed t_{av} measurements on 1 nM labelled IR-ECD, both in the absence and presence of insulin at a concentration of 100 nM, well above the expected K_{d} of the interaction. In contrast to the molecular binding measurements discussed in the previous section, all of which entailed an increase in size of the complex and therefore of t_{av} upon binding of the small, labelled ligand, here we found that t_{av} values of the fluorescent complex decreased by about 7% to 21 ±

0.1 ms compared to 22.5 ± 0.3 ms for the ligand-free species (Fig. 7A). Within our view of the underlying measurement principle, a decrease in t_{av} that occurs despite an increase in mass upon binding indicates a reduction in the overall spatial extent of the complex in the bound state. This strongly suggests that the readout is in fact dominated by significant conformational compaction of IR-ECD upon insulin binding. Note however that the X-ray structure of the *conformationally stabilized* 'apo' state and the cryo-EM structure of the liganded state suggest very similar r_{H} and D_{s} values which would entail negligible differences in t_{av} in our experiment, contrary to the experimental observations (top panel, Fig. 7A).

In order to quantitatively explore the implied conformational difference between the two states of the receptor in solution we repeated t_{av} measurements in slits of systematically smaller height in the range $h_1 \approx$ 30-60 nm (Fig. 7A, lower panel). Similar to the observations on the DNA nanostructures, we noted progressively larger disparities in t_{av} between the 'apo' and liganded states increasing from $\approx$ 5 % in the deepest slits to $>$ 20% in the shallowest (Fig. 7A). The plots in Fig. 7B graphically represent permissible values r_{H} and D_{s} for the two molecular states inferred from measurements at four different slit heights. The r_{H} and D_{s} values inferred from the intersection of lines representing a pair of measurements corresponding to mean values of h_1 =38.7 nm and 62.6 nm show that the two states are well resolved, highlighting systematic differences $\Delta D_{\mathrm{s}} = D_{\mathrm{s,lig.}} - D_{\mathrm{s,apo}} \approx 1.5$ nm and $\Delta r_{\mathrm{H}} \lesssim 0.25$ nm. Note that the uncertainty in determining the calibrated height h_1, which can be as high as ≈3-4 nm in some cases, limits the level of accuracy on absolute values of some of our present readouts of r_{H} and D_{s}, but this can be improved upon in the future (Fig. 7C). Notwithstanding the current uncertainties in determining absolute values of r_{H} and D_{s}, we note systematic measured differences between the apo and liganded states that are relatively robust to uncertainties in h_1 (Fig. 7D). Furthermore, the analysis readily admits solutions for r_{H} and D_{s} that agree reasonably with the corresponding values characterizing the cryo-EM structure of the ligand saturated state (pdb: 6SOF) for values of h_1 within the calibration error (Fig. 7C, D, black symbol).

Taken together our results indicate that the liganded state may indeed be more compact than the ligand-free 'apo' state. Whilst the absence of a 3D structure of the apo-state in solution precludes a quantitative assessment of our molecular property readouts against an independent structural model, our measurements do indeed point to an apo-state characterised by a significantly greater flexibility than the liganded-state (right panels, Fig. 7C). Thus, in contrast to the binding of the 'fluorescent bait' aptamer to insulin, the smaller t_{av} value of the IR-ECD 'bait' in the presence of insulin appears to be a direct consequence of conformational compaction as the IR-ECD transitions from the insulin-free apo to the insulin-bound liganded state. Furthermore, this conformational compaction is large enough to offset the effect of a modest increase in mass (up to $\approx$ 25 kDa) that occurs upon binding of up to 4 insulin molecules. Moreover, the relative insensitivity of the hydrodynamic radius r_{H} of IR-ECD to the presence or absence of insulin is in broad agreement with diffusion measurements on the same samples using FCS, as well with computational estimates based on the molecular 3D structures of the stabilized 'apo' and liganded states (Fig. 7C, fig. S13, table S4). Note however that our measurement does indeed detect a small increase in r_{H} of $\lesssim 4\%$ commensurate with the that expected from an increase in molecular weight. Thus, the overall measured response is dominated by the conformational change in the receptor induced by ligand binding, as reflected in the reduced diameter of the molecular bounding sphere, D_{s} (Fig. 7C, D). Our quantitative readout of molecular conformational parameters could in future be leveraged to make generalized structural inferences on 3D conformational states of molecular complexes and states in solution.

Affinity based detection of free insulin in simulated serum

Finally, Fig. 7E demonstrates a route to exploiting conformation changes in a receptor protein in order to detect analytes near physiologically relevant concentrations in complex samples. We first incubated 80 pM labelled IR-ECD with a range of insulin concentrations in PBS buffer for an hour at room temperature. Imaging the chip for about 5 min per incubated sample, we noted a gradual reduction in t_{av} with increasing insulin concentration, with the maximum and minimum timescales corresponding to those expected for the 'apo' (22.5 ms) and liganded states (21 ms) in a slit of $h_1 \approx 40$ nm. Assuming that t_{av} is proportional to the amount of IR-ECD in the liganded-state we obtain $K_{\mathrm{d}} \approx 2 - 12$ nM, in good agreement with the literature, but 1-2 orders of magnitude higher than the typical physiologically relevant concentration of free insulin (< 1 nM) (*60-64*).

We then titrated varying concentrations of insulin into simulated serum containing a fixed concentration of ≈ 100 pM of labelled IR-ECD, incubated the samples for 30 min at room temperature and assessed the binding response, as indicated by t_{av} (fig. S22). We found K_{d} values in the range of 2-4 nM, supporting detection of free insulin at concentrations of 1-10 nM in serum, which lies at the higher end of the typical physiological range. However, genetically modified and conformationally stabilized variants of IR-ECD are known to offer much higher affinities of $K_{\mathrm{d}} \approx 80$ pM, comparable with the full-length IR (*67, 68*). Given an affinity reagent with a lower K_{d} value, our current ability to quantitatively measure t_{av} at femtomolar concentrations of fluorescent species offers the sub-picomolar detection sensitivity required for highly quantitative, rapid and sensitive detection of analytes in ultra-small volumes of serum or plasma in the pM regime, based on the detection of changes in conformation and/or molecular weight (Fig. 5B, 6B, fig. S5). The same detection principle may be applied to a host of diagnostic biomarkers, including small molecule hormones and metabolites, and such detection assays may be run in a highly parallel spectrally multiplexed fashion (*69*).

Rather often, e.g., in metal-ion binding to proteins, molecular weight shifts upon binding are small compared to changes in shape and conformational properties (*70*). ETsy presents a versatile chip-based measurement platform that meets a variety of molecular analytical goals via highly size-sensitive detection under native conditions in solution. The same platform operated in "charge mode" enables state, structure and chemical discrimination based on molecular effective charge as previously described, adding a powerful dimension to the existing multiparametric molecular property readout in solution (*20, 34, 36*). Optical detection based on scattered light may deliver similar information free of covalent modification by labels, potentially enabling simultaneous molecular weight inferences (*12*).

Our current work relies on fluorescence detection, where the label serves the important function of conferring (bio)chemical specificity on the readout, which is crucial in applications focused on detecting species of interest in a complex mixture. The labelled species present at trace concentrations can act as a reporter of molecular or complex-state in the presence of high concentrations (mM) of unlabelled, interacting matter of any mass, identity and compositional complexity. Spectral and photophysical properties of labels may be further leveraged to foster optical multi-dimensionality in high throughput screening of molecular complexes (*71, 72*). Although we focus on constructing simple molecular-shape spectra containing three species in the present study, larger *N* values will support the direct extraction of a higher number of components characterised by sufficiently distinct $t_i$values (fig. S6; see supplementary text). Future improvements in the stability and longevity of the label could foster high precision measurements on just one molecule, and therefore the ability to record highly resolved molecular spectra by

permitting $N_{\mathrm{hop}}\sim 10^3 - 10^4$ per molecule in an ensemble. Finally, the ability to directly map conformation-dependent parameters detected in the measurement to corresponding molecular 3D structure-based computed quantities suggests that the ETsy readout may support general inference of molecular structure and conformation in solution, likely contributing a coarse-grained yet high-speed screening dimension to molecular structure/conformation studies. High throughput screening and identification of 3D conformation in solution may support hybrid machine learning approaches to molecular structural modelling, validation and inference.

Acknowledgments: We gratefully acknowledge Andrew Turberfield for the kind gift of DNA nanostructures, the NIH Tetramer Core Facility (contract number 75N93020D00005) for providing HLA-A*03:01, and Robin W. Klemm for helpful discussions.

Funding:

European Research Council Consolidator Grant 724180 (XZ, TJDB, MK)

Biotechnology and Biological Sciences Research Council BB/W017415/1 (DS, XZ, MK)

Wellcome LEAP Delta Tissue Program (XZ, KCZ, JLPB, MK)

Royal Society Research Grants RGS\R2\242523 (AHES)

Royal Society of Chemistry Research grant R24-8767374007 (AHES)

Federal Ministry of Education and Research (BMBF) and the German Center for Diabetes Research (DZD e.V.) (UC)

Author contributions:

Conceptualization and model: MK
Methodology: XZ, TJDB, MK
Measurement: XZ, TJDB, DS
Data analysis and processing: TJDB
Protein synthesis/recombinant production, labelling and quality control: KCZ, MG, JLPB, UC
Nucleic acid design, synthesis and quality control: AHES
Visualization: TJDB, XZ, KCZ, DS, MK
Supervision: MK
Writing – original draft: MK
Manuscript – review & editing: All authors

Competing interests: XZ, TJDB and KCZ are contributors on an intellectual property filing by MK.

Data and materials availability: All data are available in the main text or the supplementary materials.

Supplementary Materials

Materials and Methods
Supplementary Text
Figs. S1 to S22
Tables S1 to S5
References (*73–96*)
Movies S1 to S2

References

1. Q. Wang, W. E. Moerner, Single-molecule motions enable direct visualization of biomolecular interactions in solution. *Nature Methods* **11**, 555–558 (2014).
2. H. Wilson, Q. Wang, ABEL-FRET: tether-free single-molecule FRET with hydrodynamic profiling. *Nature Methods* **18**, 816–820 (2021).
3. J. A. Jones, D. K. Wilkins, L. J. Smith, C. M. Dobson, Characterisation of protein unfolding by NMR diffusion measurements. *Journal of Biomolecular NMR* **10**, 199–203 (1997).
4. D. K. Wilkins *et al.*, Hydrodynamic Radii of Native and Denatured Proteins Measured by Pulse Field Gradient NMR Techniques. *Biochemistry* **38**, 16424-16431 (1999).
5. R. Russell, I. S. Millett, S. Doniach, D. Herschlag, Small angle X-ray scattering reveals a compact intermediate in RNA folding. *Nature Structural Biology* **7**, 367-370 (2000).
6. J. E. Kohn *et al.*, Random-coil behavior and the dimensions of chemically unfolded proteins. *Proceedings of the National Academy of Sciences* **101**, 12491-12496 (2004).
7. S. B. Dubin, J. H. Lunacek, G. B. Benedek, Observation of the spectrum of light scattered by solutions of biological macromolecules. *Proceedings of the National Academy of Sciences* **57**, 1164 - 1171 (1967).
8. M. Weissman, H. Schindler, G. Feher, Determination of molecular weights by fluctuation spectroscopy: application to DNA. *Proceedings of the National Academy of Sciences* **73**, 2776-2780 (1976).
9. T. Ha *et al.*, Single-molecule fluorescence spectroscopy of enzyme conformational dynamics and cleavage mechanism. *Proceedings of the National Academy of Sciences* **96**, 893-898 (1999).
10. C. B. Müller *et al.*, Precise measurement of diffusion by multi-color dual-focus fluorescence correlation spectroscopy. *Europhysics Letters* **83**, 46001 (2008).
11. G. Young *et al.*, Quantitative mass imaging of single biological macromolecules. *Science* **360**, 423-427 (2018).
12. B. Špačková *et al.*, Label-free nanofluidic scattering microscopy of size and mass of single diffusing molecules and nanoparticles. *Nature Methods* **19**, 751-758 (2022).
13. M. Dahmardeh *et al.*, Self-supervised machine learning pushes the sensitivity limit in label-free detection of single proteins below 10 kDa. *Nature Methods* **20**, 442–447 (2023).
14. L.-M. Needham *et al.*, Label-free detection and profiling of individual solution-phase molecules. *Nature* **629**, 1062–1068 (2024).
15. S. Ohayon *et al.*, Full-Length Single Protein Molecules Tracking and Counting in Thin Silicon Channels. *Advanced Materials* **36**, 2314319 (2024).
16. J. Han, H. G. Craighead, Separation of Long DNA Molecules in a Microfabricated Entropic Trap Array. *Science* **288**, 1026-1029 (2000).
17. J. Fu, R. B. Schoch, A. L. Stevens, S. R. Tannenbaum, J. Han, A patterned anisotropic nanofluidic sieving structure for continuous-flow separation of DNA and proteins. *Nature Nanotechnology* **2**, 121–128 (2007).
18. E. V. Yates *et al.*, Latent analysis of unmodified biomolecules and their complexes in solution with attomole detection sensitivity. *Nature Chemistry* **7**, 802–809 (2015).
19. E. C. Yusko *et al.*, Real-time shape approximation and fingerprinting of single proteins using a nanopore. *Nature Nanotechnology* **12**, 360–367 (2016).
20. F. Ruggeri *et al.*, Single-molecule electrometry. *Nature Nanotechnology* **12**, 488-495 (2017).

21. F. Ruggeri, M. Krishnan, Entropic Trapping of a Singly Charged Molecule in Solution. *Nano Letters* **18**, (2018).
22. S. Schmid, P. Stömmer, H. Dietz, C. Dekker, Nanopore electro-osmotic trap for the label-free study of single proteins and their conformations. *Nature Nanotechnology* **16**, 1244–1250 (2021).
23. G. Krainer *et al.*, Direct digital sensing of protein biomarkers in solution. *Nature Communications* **14**, 653 (2023).
24. R. P. B. Jacquat *et al.*, Single-Molecule Sizing through Nanocavity Confinement. *Nano Letters* **23**, 1629-1636 (2023).
25. I. I. Hosseini *et al.*, Tunable nanofluidic device for digital nucleic acid analysis. *Nanoscale* **16**, 9583-9592 (2024).
26. G. Krainer *et al.*, Single-molecule digital sizing of proteins in solution. *Nature Communications* **15**, 7740 (2024).
27. M. Friedel, A. Baumketner, J.-E. Shea, Effects of surface tethering on protein folding mechanisms. *Proceedings of the National Academy of Sciences* **103**, 8396-8401 (2006).
28. Y. Pang, R. Gordon, Optical Trapping of a Single Protein. *Nano Letters* **12**, 402-406 (2012).
29. F. Perrin, Mouvement Brownien d'un ellipsoide (II). Rotation libre et dépolarisation des fluorescences. Translation et diffusion de molécules ellipsoidales. *Journal de Physique et le Radium* **7**, 1-11 (1936).
30. A. Ortega, D. Amorós, J. García de la Torre, Prediction of Hydrodynamic and Other Solution Properties of Rigid Proteins from Atomic- and Residue-Level Models. *Biophysical Journal* **101**, 892-898 (2011).
31. P. J. Fleming, K. G. Fleming, HullRad: Fast Calculations of Folded and Disordered Protein and Nucleic Acid Hydrodynamic Properties. *Biophysical Journal* **114**, 856-869 (2018).
32. E. Haustein, P. Schwille, Fluorescence Correlation Spectroscopy: Novel Variations of an Established Technique. *Annual Review of Biophysics* **36**, 151-169 (2007).
33. A. G. Kikhney, D. I. Svergun, A practical guide to small angle X-ray scattering (SAXS) of flexible and intrinsically disordered proteins. *FEBS Letters* **589**, 2570-2577 (2015).
34. M. Bespalova, A. Behjatian, N. Karedla, R. Walker-Gibbons, M. Krishnan, Opto-Electrostatic Determination of Nucleic Acid Double-Helix Dimensions and the Structure of the Molecule–Solvent Interface. *Macromolecules* **55**, 6200-6210 (2022).
35. R. Walker-Gibbons, X. Zhu, A. Behjatian, T. J. D. Bennett, M. Krishnan, Sensing the structural and conformational properties of single-stranded nucleic acids using electrometry and molecular simulations. *Scientific Reports* **14**, 20582 (2024).
36. G. Kloes *et al.*, Far-Field Electrostatic Signatures of Macromolecular 3D Conformation. *Nano Letters* **22**, 7834–7840 (2022).
37. P. J. Flory, *Principles of Polymer Chemistry*. (Cornell University Press, 1953).
38. Y. Nojima, K. Iguchi, Y. Suzuki, A. Sato, The pH-Dependent Formation of PEGylated Bovine Lactoferrin by Branched Polyethylene Glycol (PEG)-N-Hydroxysuccinimide (NHS) Active Esters. *Biological and Pharmaceutical Bulletin* **32**, 523-526 (2009).
39. F. Ruggeri, M. Krishnan, Lattice diffusion of a single molecule in solution. *Physical Review E* **96**, 062406 (2017).
40. R. E. Dickerson *et al.*, The Anatomy of A-, B-, and Z-DNA. *Science* **216**, 475-485 (1982).
41. A. Wynveen, D. J. Lee, A. A. Kornyshev, S. Leikin, Helical coherence of DNA in crystals and solution. *Nucleic Acids Research* **36**, 5540–5551 (2008).

42. R. P. Goodman *et al.*, Rapid Chiral Assembly of Rigid DNA Building Blocks for Molecular Nanofabrication. *Science* **310**, 1661-1665 (2005).
43. S. Fischer *et al.*, Shape and Interhelical Spacing of DNA Origami Nanostructures Studied by Small-Angle X-ray Scattering. *Nano Letters* **16**, 4282-4287 (2016).
44. E. Poppleton, R. Romero, A. Mallya, L. Rovigatti, P. Sulc, OxDNA.org: a public webserver for coarse-grained simulations of DNA and RNA nanostructures. *Nucleic Acids Research* **49**, W491-W498 (2021).
45. A. A. Istratov, O. F. Vyvenko, Exponential analysis in physical phenomena. *Review of Scientific Instruments* **70**, 1233–1257 (1999).
46. R. J. L. Townshend *et al.*, Geometric deep learning of RNA structure. *Science* **373**, 1047-1051 (2021).
47. K. Zhang *et al.*, Cryo-EM structure of a 40 kDa SAM-IV riboswitch RNA at 3.7 Å resolution. *Nature Communications* **10**, 5511 (2019).
48. A. D. Garst, A. Héroux, R. P. Rambo, R. T. Batey, Crystal Structure of the Lysine Riboswitch Regulatory mRNA Element. *Journal of Biological Chemistry* **283**, 22347-22351 (2008).
49. C. D. Stoddard *et al.*, Free State Conformational Sampling of the SAM-I Riboswitch Aptamer Domain. *Structure* **18**, 787-797 (2010).
50. W. Palau, C. Di Primo, Single-cycle kinetic analysis of ternary DNA complexes by surface plasmon resonance on a decaying surface. *Biochimie* **94**, 1891-1899 (2012).
51. I. Jarmoskaite, I. AlSadhan, P. P. Vaidyanathan, D. Herschlag, How to measure and evaluate binding affinities. *eLife* **9**, e57264 (2020).
52. J. SantaLucia, D. Hicks, The thermodynamics of DNA structural motifs. *Annu Rev Bioph Biom* **33**, 415-440 (2004).
53. H. Wilson, M. Lee, Q. Wang, Probing DNA-protein interactions using single-molecule diffusivity contrast. *Biophysical Reports* **1**, 100009 (2021).
54. C. J. Barnstable *et al.*, Production of monoclonal antibodies to group A erythrocytes, HLA and other human cell surface antigens-new tools for genetic analysis. *Cell* **14**, 9-20 (1978).
55. M. N. Hug *et al.*, HLA antibody affinity determination: From HLA-specific monoclonal antibodies to donor HLA specific antibodies (DSA) in patient serum. *HLA* **102**, 278-300 (2023).
56. M. M. Schneider *et al.*, Microfluidic antibody affinity profiling of alloantibody-HLA interactions in human serum. *Biosensors and Bioelectronics* **228**, 115196 (2023).
57. W. Yoshida *et al.*, Selection of DNA aptamers against insulin and construction of an aptameric enzyme subunit for insulin sensing. *Biosensors and Bioelectronics* **24**, 1116-1120 (2009).
58. Z. Hao *et al.*, Real-Time Monitoring of Insulin Using a Graphene Field-Effect Transistor Aptameric Nanosensor. *ACS Applied Materials & Interfaces* **9**, 27504–27511 (2017).
59. D. Wu *et al.*, Flow-Cell-Based Technology for Massively Parallel Characterization of Base-Modified DNA Aptamers. *Analytical Chemistry* **95**, 2645–2652 (2023).
60. J. G. Menting *et al.*, How insulin engages its primary binding site on the insulin receptor. *Nature* **493**, 241-245 (2013).
61. T. Gutmann *et al.*, Cryo-EM structure of the complete and ligand-saturated insulin receptor ectodomain. *Journal of Cell Biology* **219**, e201907210 (2020).
62. P. De Meyts, J. Whittaker, Structural biology of insulin and IGF1 receptors: implications for drug design. *Nature Reviews Drug Discovery* **1**, 769-783 (2002).

63. V. V. Kiselyov, S. Versteyhe, L. Gauguin, P. D. Meyts, Harmonic oscillator model of the insulin and IGF1 receptors' allosteric binding and activation. *Molecular Systems Biology* **5**, 243 (2009).
64. T. Gutmann, K. H. Kim, M. Grzybek, T. Walz, Ü. Coskun, Visualization of ligand-induced transmembrane signaling in the full-length human insulin receptor. *Journal of Cell Biology* **217**, 1643–1649 (2018).
65. I. Croll, Tristan *et al.*, Higher-Resolution Structure of the Human Insulin Receptor Ectodomain: Multi-Modal Inclusion of the Insert Domain. *Structure* **24**, 469-476 (2016).
66. X. Xiong *et al.*, Symmetric and asymmetric receptor conformation continuum induced by a new insulin. *Nature Chemical Biology* **18**, 511-519 (2022).
67. P. A. Hoyne *et al.*, High affinity insulin binding by soluble insulin receptor extracellular domain fused to a leucine zipper. *FEBS Letters* **479**, 15-18 (2000).
68. F. Weis *et al.*, The signalling conformation of the insulin receptor ectodomain. *Nature Communications* **9**, 4420 (2018).
69. F. Ruggeri, M. Krishnan, Spectrally resolved single-molecule electrometry. *The Journal of Chemical Physics* **148**, 123307 (2018).
70. F. Kilár, I. Simon, The effect of iron binding on the conformation of transferrin. A small angle x-ray scattering study. *Biophysical Journal* **48**, 799-802 (1985).
71. H. Hwang, S. Myong, Protein induced fluorescence enhancement (PIFE) for probing protein–nucleic acid interactions. *Chemical Society Reviews* **43**, 1221-1229 (2014).
72. A. Langer *et al.*, A New Spectral Shift-Based Method to Characterize Molecular Interactions. *ASSAY and Drug Development Technologies* **20**, 83-94 (2022).
73. M. T. Marty *et al.*, Bayesian Deconvolution of Mass and Ion Mobility Spectra: From Binary Interactions to Polydisperse Ensembles. *Analytical Chemistry* **87**, 4370–4376 (2015).
74. M. Krishnan, N. Mojarad, P. Kukura, V. Sandoghdar, Geometry-induced electrostatic trapping of nanometric objects in a fluid. *Nature* **467**, 692-695 (2010).
75. J. R. Taylor, *An introduction to error analysis : the study of uncertainties in physical measurements*. (University Science Books, Sausalito, Calif, ed. 2nd, 1997).
76. G. Emilsson *et al.*, Strongly Stretched Protein Resistant Poly(ethylene glycol) Brushes Prepared by Grafting-To. *ACS Applied Materials & Interfaces* **7**, 7505-7515 (2015).
77. Boussinesq, Mémoire sur l'influence des frottements dans les mouvements réguliers des fluides. *Journal de Mathématiques Pures et Appliquées* **13**, 377-424 (1868).
78. S. H. Banyard, D. K. Stammers, P. M. Harrison, Electron density map of apoferritin at 2.8-Å resolution. *Nature* **271**, 282-284 (1978).
79. T. Dertinger *et al.*, Two-Focus Fluorescence Correlation Spectroscopy: A New Tool for Accurate and Absolute Diffusion Measurements. *ChemPhysChem* **8**, 433-443 (2007).
80. S. Broersma, Viscous Force Constant for a Closed Cylinder. *The Journal of Chemical Physics* **32**, 1632–1635 (1960).
81. M. M. Tirado, J. G. De La Torre, Translational friction coefficients of rigid, symmetric top macromolecules. Application to circular cylinders. *The Journal of Chemical Physics* **71**, 2581-2587 (1979).
82. J. G. De La Torre, V. A. Bloomfield, Hydrodynamic properties of complex, rigid, biological macromolecules: theory and applications. *Quarterly Reviews of Biophysics* **14**, 81-139 (1981).
83. H. C. Berg, *Random Walks in Biology*. (Princeton University Press, 1984).
84. A. Ortega, J. García de la Torre, Hydrodynamic properties of rodlike and disklike particles in dilute solution. *The Journal of Chemical Physics* **119**, 9914–9919 (2003).

85. S. H. Koenig, Brownian motion of an ellipsoid. A correction to Perrin's results. *Biopolymers* **14**, 149-160 (1975).
86. P. V. Ruijgrok, N. R. Verhart, P. Zijlstra, A. L. Tchebotareva, M. Orrit, Brownian Fluctuations and Heating of an Optically Aligned Gold Nanorod. *Physical Review Letters* **107**, 037401 (2011).
87. T. E. Ouldridge, A. A. Louis, J. P. K. Doye, Structural, mechanical, and thermodynamic properties of a coarse-grained DNA model. *Journal of Chemical Physics* **134**, 085101 (2011).
88. E. Poppleton *et al.*, Design, optimization and analysis of large DNA and RNA nanostructures through interactive visualization, editing and molecular simulation. *Nucleic Acids Research* **48**, e72 (2020).
89. F. E. Arrighi, M. Mandel, J. Bergendahl, T. C. Hsu, Buoyant densities of DNA of mammals. *Biochemical Genetics* **4**, 367-376 (1970).
90. S. R. De Kloet, B. A. G. Andrean, Buoyant density gradient centrifugation of RNA and DNA in alkali iodide solutions. *Biochimica et Biophysica Acta (BBA) - Nucleic Acids and Protein Synthesis* **247**, 519-527 (1971).
91. H. Cramér, *Mathematical Methods of Statistics*. Goldstine Printed Materials (Princeton University Press, 1946).
92. M. Zuker, Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Research* **31**, 3406-3415 (2003).
93. J. D. Altman *et al.*, Phenotypic Analysis of Antigen-Specific T Lymphocytes. *Science* **274**, 94-96 (1996).
94. S. S. Chandran *et al.*, Immunogenicity and therapeutic targeting of a public neoantigen derived from mutated PIK3CA. *Nature Medicine* **28**, 946-957 (2022).
95. W. Yoshida *et al.*, Selection of DNA aptamers against insulin and construction of an aptameric enzyme subunit for insulin sensing. *Biosensors and Bioelectronics* **24**, 1116-1120 (2009).
96. Y. Wu, B. Midinov, R. J. White, Electrochemical Aptamer-Based Sensor for Real-Time Monitoring of Insulin. *ACS Sensors* **4**, 498-503 (2019).

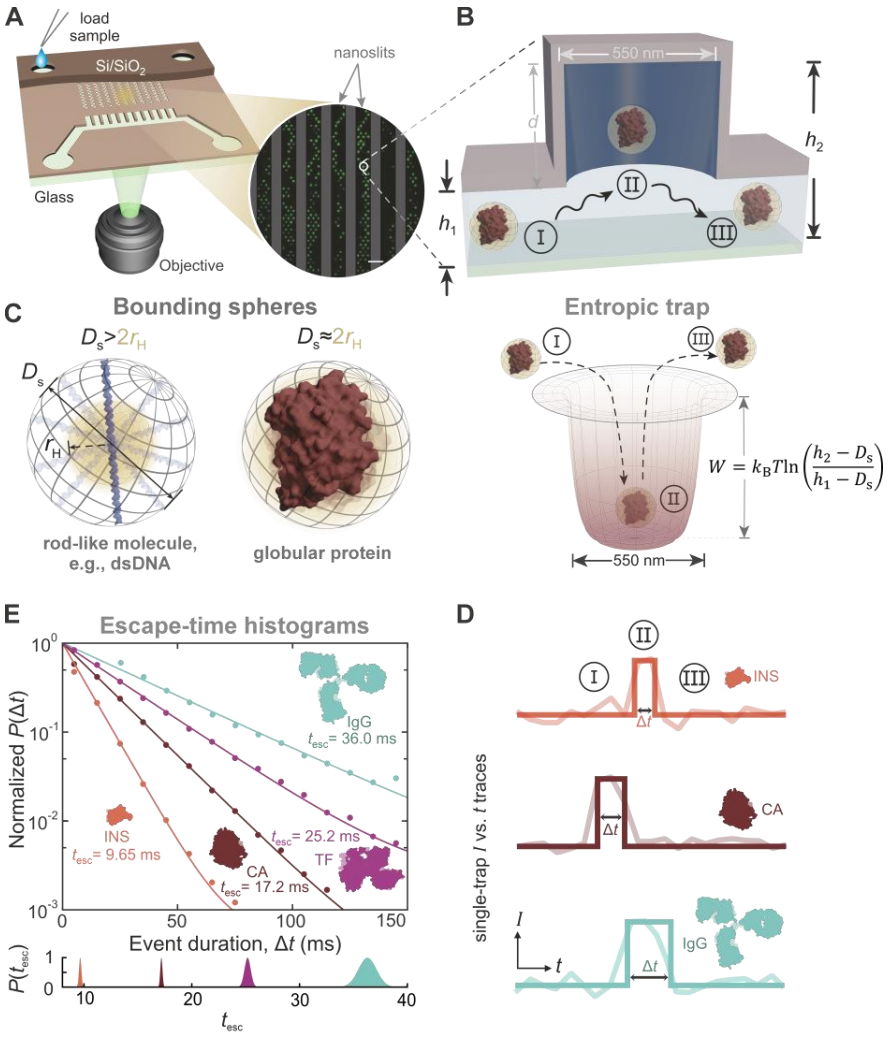


Fig. 1. Principle and overview of escape-time molecular size-and-shape spectrometry (ETsy) in solution. (**A**) Schematic representation of a microfluidic chip carrying a series of fluid-filled parallel-plate nanoslits in which arrays of geometric indentations create local entropic traps for freely diffusing fluorescently labelled molecules in solution. Wide-field fluorescence microscopy image presents the central quarter of a typical field of view, time-averaged over 1 s, displaying ~25 molecules diffusing in the trap landscape. Scale bar 5 μm. (**B**) Schematic depictions of a globular protein molecule entering (state-I), residing in (state-II) and leaving (state-III) a nanostructured indentation of diameter 550 nm, depth $d \approx 220 - 290$ nm and height $h_2 = h_1 + d$ nm located in a nanoslit of height $h_1 = 20 - 70$ nm (top). Bottom - corresponding entropic potential well of depth W in which molecules reside for an average time given approximately by $t_{\mathrm{esc}} \propto r_{\mathrm{H}} \exp(W/k_{\mathrm{B}}T)$ where r_{H} is the hydrodynamic radius, k_{B} is the Boltzmann constant and T is the absolute temperature. (**C**) Graphic representation of bounding spheres (grey) of diameter D_{s} enclosing a rod-like double-stranded (ds) DNA molecule (left) and a globular protein molecule (right). r_{H} (yellow sphere) can be considerably smaller than $D_{\mathrm{s}}/2$ for a non-spherical molecule, in contrast to a globular molecule. D_{s} and r_{H} values for a given species maybe computed from a molecular structural model as described in the supplementary text. (**D**) Intensity vs. time trace of each trap locus in the lattice permits identification of molecular residence or escape events of duration Δt. Representative events detected using a step-finding algorithm (dark lines)

superimposed on raw traces. Three typical escape-time events for insulin (INS), carbonic anhydrase (CA), and immunoglobulin G (IgG) illustrating that Δt depends on the molecular weight of the species. (**E**) Probability density distributions $P(\Delta t)$ vs. Δt for $N = 10^4 - 10^6$ recorded escape events fit to the form $P(\Delta t) \propto \exp(-\Delta t/t_{\mathrm{esc}})$ to extract means t_{esc} values. Gaussian probability distributions, $P(t_{\mathrm{esc}})$, centered on the mean fit value t_{esc} with a standard deviation given by the corresponding fit error (or standard error of the mean) are displayed for each measurement and denote the probability distribution of mean t_{esc} values expected in multiple realisations of the measurement (see supplementary text).

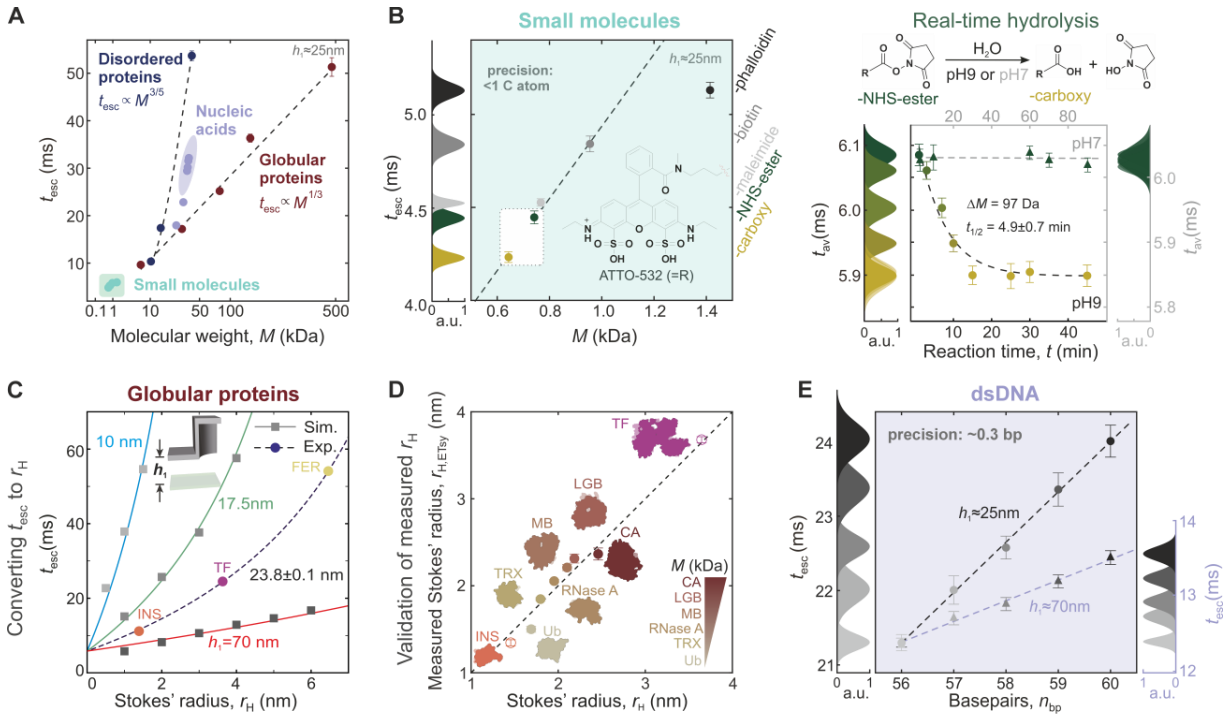


Fig. 2. Sensitivity of measured escape-times to molecular weight. (**A**) Measurements of escape time t_{esc} vs. Molecular weight, M, over approximately 3 orders of magnitude in molecular weight including small organic molecules (green symbols), globular proteins (brown symbols), disordered proteins (blue symbols) and dsDNA (purple symbols) in a device with $h_1 \approx 25$ nm. Globular proteins (insulin, carbonic anhydrase, transferrin, Anti-insulin IgG, and apoferritin) and disordered proteins (*20*) (Starmaker-like - Stm-l, Prothymosin α - ProTα, and a 13-mer proline-rich polypeptide sequence given in table S2) reveal good agreement with scaling behaviours $t_{esc} \propto M^{1/3}$ and $M^{3/5}$ respectively. (**B**) Plot of t_{esc} vs. M for small fluorescent organic molecules represented by different chemical derivatives of the fluorophore ATTO 532 (left). Approximately 0.5% measurement precision on t_{esc} implies ability to detect differences of ≈ 10 Da in the 1 kDa molecular weight regime. Right - Real-time tracking of the hydrolysis of the NHS ester derivative into the carboxyl form entails a change in molecular weight of 97 Da. The 3% overall change in t_{av} over the course of the reaction is much larger than the measurement precision of ≈0.5% in a minute-long measurement window permitting rate of hydrolysis to be measured at pH 9 where the reaction has a half-life of ca. 5 min (details in supplementary text). (**C**) Measured t_{esc} values may be converted to r_H and D_s values describing the molecular species using BD simulations (as described in supplementary text). Simulated values for spheres of radius r_H (square symbols) in a device with $d = 300$ nm illustrate the increasing non-linearity of the measured response with decreasing h_1, and are well fit by the calibration function Eq. (1) with $A \approx 0.3$ s/μm and $t_o = 5.83$ ms which is used for all protein experiments (red, green and blue curves). An example calibration measurement to determine the value of h_1 of the experimental device is performed using insulin (orange, INS), transferrin (purple, TF) and apoferritin (yellow, FER) whose Stokes' radii, r_H, are taken from molecular structures (as outlined in supplementary text). Measured t_{esc} data are fitted with the calibration function yielding $h_1 = 23.8 \pm 0.1$ nm which agrees with the nominal design

value of $h_1 \approx 25$ nm (dashed curve). (**D**) r_{H} values measured for a test set of proteins in the range $M = 10 - 30$ kDa (filled symbols: ubiquitin - Ub, thioredoxin - TRX, RNase A, myoglobin - MB, β-lactoglobulin – LGB, carbonic anhydrase – CA; open symbols – INS and TF calibrator molecules) compare well with theoretical expectations, as well as with measurements using fluorescence correlation spectroscopy (FCS) (see fig. S9, table S4 and supplementary text). (**E**) Single base-pair differences in dsDNA in the range of 56-60 bp are clearly distinguishable, with the magnitude of measured differences increasing strongly with decreasing h_1.

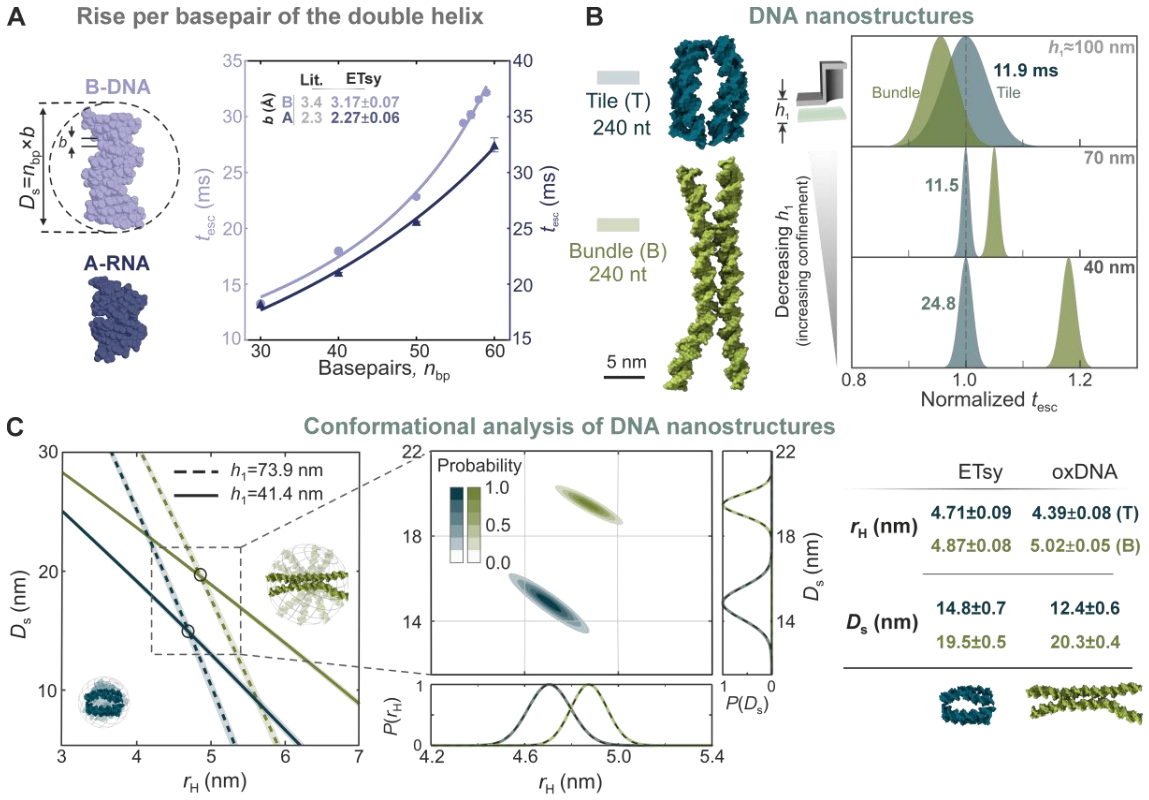


Fig. 3. Sensitivity of the escape-time readout to molecular shape (3D conformation). (**A**) Measurements of escape time t_{esc} vs. number of basepairs, n_{bp}, for a series of dsDNA (B-form double helix) and dsRNA (A-form helix), fitted with Eq. (1) yielding $A = 0.35 \pm 0.01$ s/μm for RNA, and 0.18 ± 0.01 s/μm for DNA, and rise per basepair values, b, that are in good agreement with literature values from high resolution structural methods (*40*). (**B**) DNA nanostructures composed of an identical number of nucleotides (nt) but in different 3D conformations of a 'square tile' and 'bundle' display increasingly separated measured t_{esc} values with decreasing slight height h_1. (**C**) Graphical representation of the solution of Eq. (1) using t_{esc} measurements for two different mean calibrated values of h_1 (solid and dashed lines, see supplementary text for details). Shaded bands around lines reflect the precision of the t_{esc} measurement. Region of intersection of the two uncertainty broadened lines yields estimates of D_{s} and r_{H} for each nanostructure (left) and Gaussian probability density distributions of molecular physical properties $P(D_{\mathrm{s}}, r_{\mathrm{H}})$, $P(r_{\mathrm{H}})$ and $P(D_{\mathrm{s}})$ (middle panel). Dashed superimposed curves demonstrate that the widths of the distributions reflect the statistical uncertainty on t_{esc} measurement and do not stem from detected conformational flexibility of the nanostructures. Table comparing ETsy measurements with inferences from coarse-grained structures simulated using oxDNA (right, see supplementary text for details).

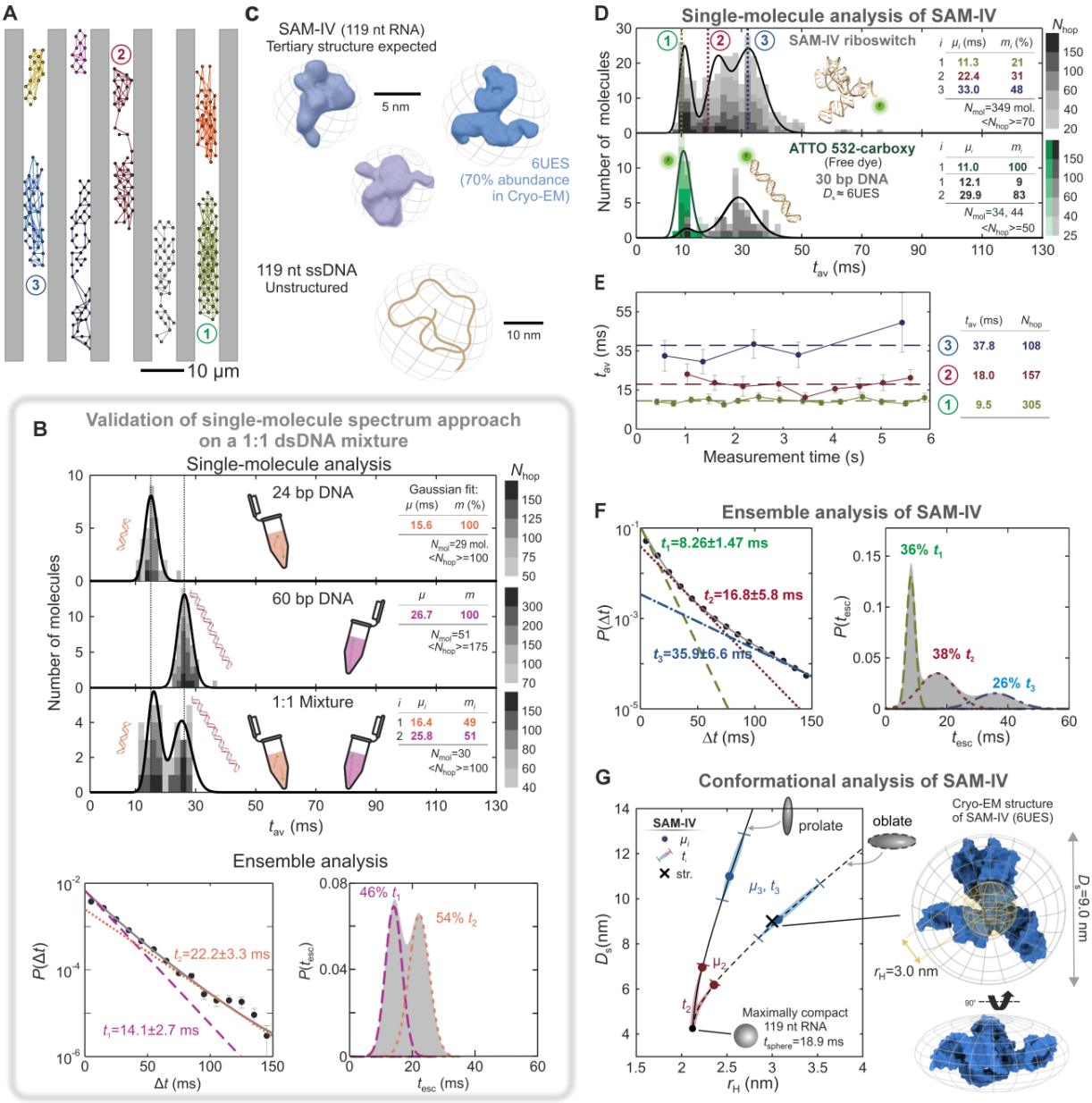


Fig. 4. Single molecule spectra of SAM-IV riboswitch conformation and comparison with a structural model. (**A**) Schematic depiction of measured individual molecular trajectories of the SAM-IV riboswitch diffusing in a landscape of entropic traps in slits of height $h_1 \approx 20$ nm (left). (**B**) Validation of the single-molecule escape-time spectrum approach on a 1:1 mixture of 24 bp and 60 DNA (top). Each integer entry in the histogram corresponds to the average escape time, t_{av}, recorded for one molecular trajectory, with the grey level denoting the observed number of hops, N_{hop}. Pure 24 bp and 60 bp DNA samples show single peaks of average times, $\mu \approx 15$ and 26 ms, whilst the mixture displays 2 components of approximately equal fractional weight, m_i. Fits of the spectra are performed according to Eq. (S24) assuming an average number of hops $\langle N_{\mathrm{hop}} \rangle$ as listed (black curve, see supplementary text for detail). The same single molecule

trajectory data pooled and fit with a bi-exponential reveals two timescales and fractional weights that agree within error with the single molecule trajectory analysis (bottom). (**C**) Multiple possible 3D conformations reported for the SAM-IV riboswitch compared with a 119 nt unstructured ssDNA (see fig. S12). (**D**) Single-molecule spectra of the SAM-IV riboswitch (top), accompanied by superimposed spectra of ATTO 532-carboxy dye and 30 bp DNA (bottom). Fitted timescales and component fractions μ_i and m_i in each case permit assignment of species and molecular states to the 3 inferred peaks in the ETsy spectrum of SAM-IV. State-1 maps on to free dye, also present in other nucleic acid samples ($\mu_1 \approx 12$ ms). Two distinct conformations attributable to SAM-IV (States 2 and 3) are much more compact than the unstructured 119 nt ssDNA control (see fig. S16), and display μ_i values similar to 30 dsDNA. (**E**) Recorded time traces of sequential escape events Δt, averaged over 16 consecutive escape events, for molecules 1, 2 and 3, display distinct and stable overall average values t_{av} over the > 6 s measurement period. (**F**) Ensemble measurements and multi-exponential fit of $N \approx 10^5$ hops of SAM-IV reveal 3 timescales, t_i, and weight fractions that capture the detected single molecule spectral timescales and compositions. (**G**) Theoretical $(D_{\mathrm{s}}, r_{\mathrm{H}})$ plot for objects of fixed volume 40 nm^3 – corresponding to an RNA molecule with $M = 38.51$ kDa characterizing SAM-IV – presents limiting cases for oblate (dashed line) and prolate ellipsoids (solid line) which intersect at a point given by the most compact, spherical state corresponding to an escape time $t_{\mathrm{sphere}} = 18.9$ ms (see supplementary text for details). Measured escape times and associated uncertainties in (**D**) and (**F**) imply $(D_{\mathrm{s}}, r_{\mathrm{H}})$ values as follows: values from single-molecule measurement μ_3(blue symbol) and the mean ensemble-fit timescale t_3 (blue shaded zone) indicate $(D_{\mathrm{s}}, r_{\mathrm{H}})$ coordinates describing a structure centered on either the 'oblate' or 'prolate' lines. Inferred coordinates for an oblate ellipsoid lie very close to the point characterizing the 6UES structure (black cross). Timescale μ_2 (red symbol) is consistent with a more compact state, on either the 'oblate' or 'prolate' lines, with red shaded regions corresponding to indications from ensemble-measurement value t_2.

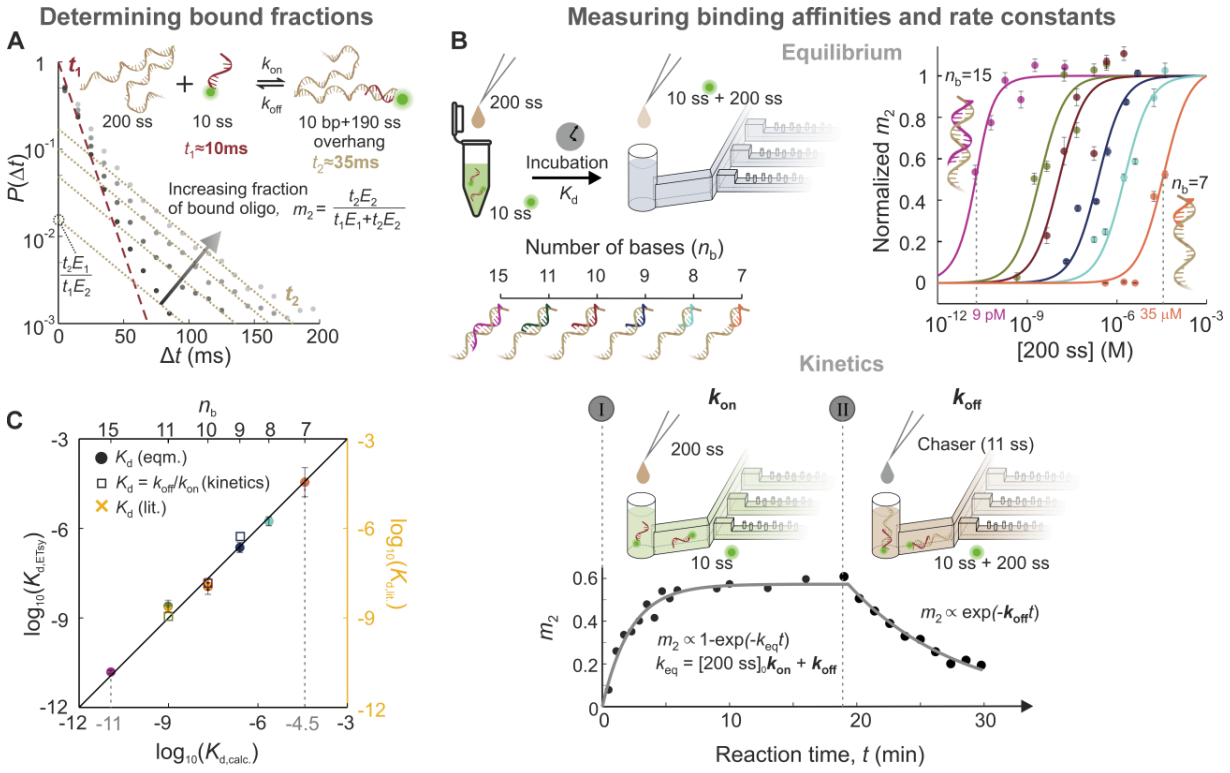


Fig. 5. Measuring association and dissociation rates, and binding affinities in DNA hybridization. (**A**) Illustration of the experimental approach to determine the fractions of free and bound molecules in a bi-molecular binding reaction. Here, a labelled 10 ssDNA oligo binds to an unlabelled 200 ssDNA fragment. Incubation with a high concentration of 200 ssDNA shifts the free oligo timescale from $t_1 \approx 12$ ms to the fully-bound timescale $t_2 \approx 21$ ms. At intermediate concentrations, escape time histograms obtained in 1-min long measurements are fit with a bi-exponential holding t_1 and t_2 fixed to yield amplitudes of two exponentials reflecting the bound fractions m_1 and m_2. E_i denote event fractions which are distinct from fractional species abundances, m_i (see supplementary text.). (**B**) Measurements of m_2 following incubation of 0.02-1 nM of labelled oligo with 200 ssDNA of increasing concentration are fit with Eq. (S29) to obtain the affinity, K_d, of the interaction. Measurements for a range of oligos of length $n_b = 7$ to 15 bases demonstrates K_d measurements in the range 10^{-11} M to 10^{-4} M (top right). Constructing $P(\Delta t)$ histograms of the reaction mixture in real time following mixing provides a measure of m_2 vs. reaction time, t, which is fit with Eq. (S30) to obtain the effective on-rate constant k_{eq} (bottom left). Equilibrated, fully bound mixtures of oligo and 200 ssDNA are mixed with an excess of unlabelled oligo, and the decaying bound fraction, m_2, measured as a function of time to yield the dissociation rate constant k_{off} using Eq. (S32) (bottom right). k_{eq} and k_{off} together determine the association rate constant, k_{on}, from which K_d can be determined independently using Eq. (S33). (**C**) Measurements of K_d from both equilibrium (circles) and kinetic (squares) approaches compare well with theoretically estimated and literature values (crosses) over a range of nearly 7 orders of magnitude (see supplementary text for further discussion).

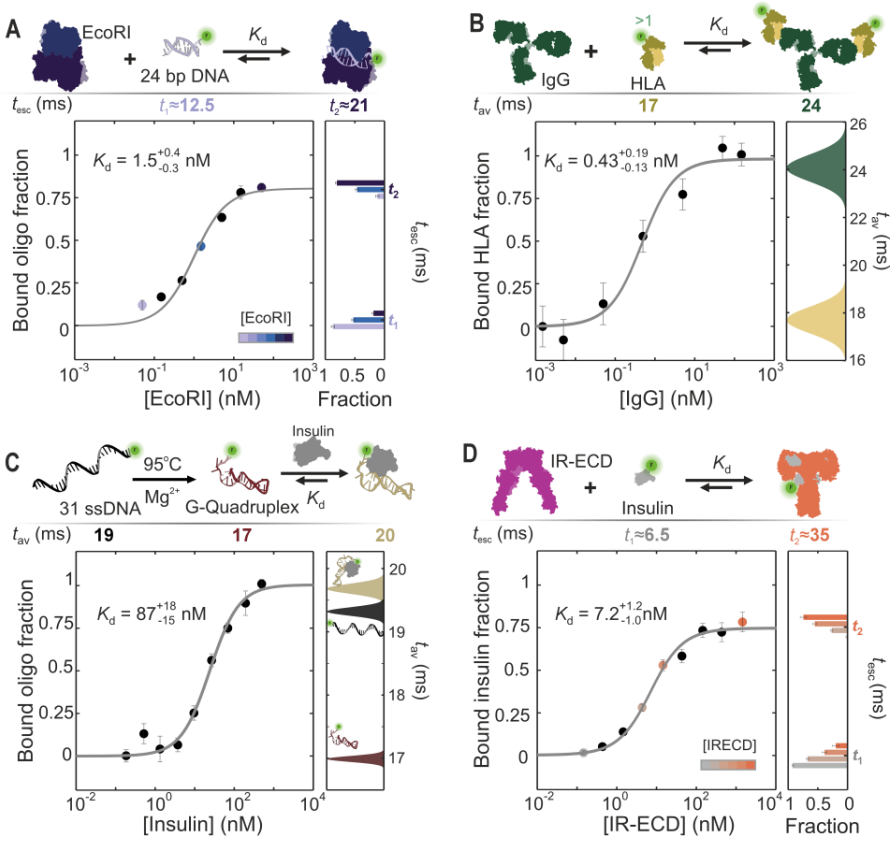


Fig. 6. Measuring binding affinities in DNA-protein and protein-protein interactions. When disparities in escape times characterizing the bound and free states, t_1 and t_2, are larger than a factor 2 the method in Fig. 5A can be used to determine m_2 and therefore K_d values for the interaction of (**A**) the restriction enzyme EcoRI (MW ca. 28.5 kDa) with 250 pM of a labelled 24 bp DNA oligo (ca. 15 kDa) carrying the restriction sequence, (**D**) the insulin receptor ectodomain (IR-ECD – 280 kDa) with 2 nM of fluorescently labelled insulin (ca. 6 kDa). When the shift in timescale upon binding is much smaller (ca. 5-10%) we use the average escape time, t_av as an indicator of m_2 (see fig. S19 and supplementary text). Using this method, we determine interaction affinities of (**B**) 100 pM of the fluorescently labelled HLA antigen (ca. 48 kDa) with the corresponding Immunoglobulin G (IgG – ca. 150 kDa). (**C**) Folding of the insulin-binding 31-base IGA3 aptamer results in a 10% reduction of t_av that subsequently systematically increases upon binding of increasing concentrations of insulin added to 5 nM of the aptamer. In all cases, the measured K_d values are in good agreement with the literature.

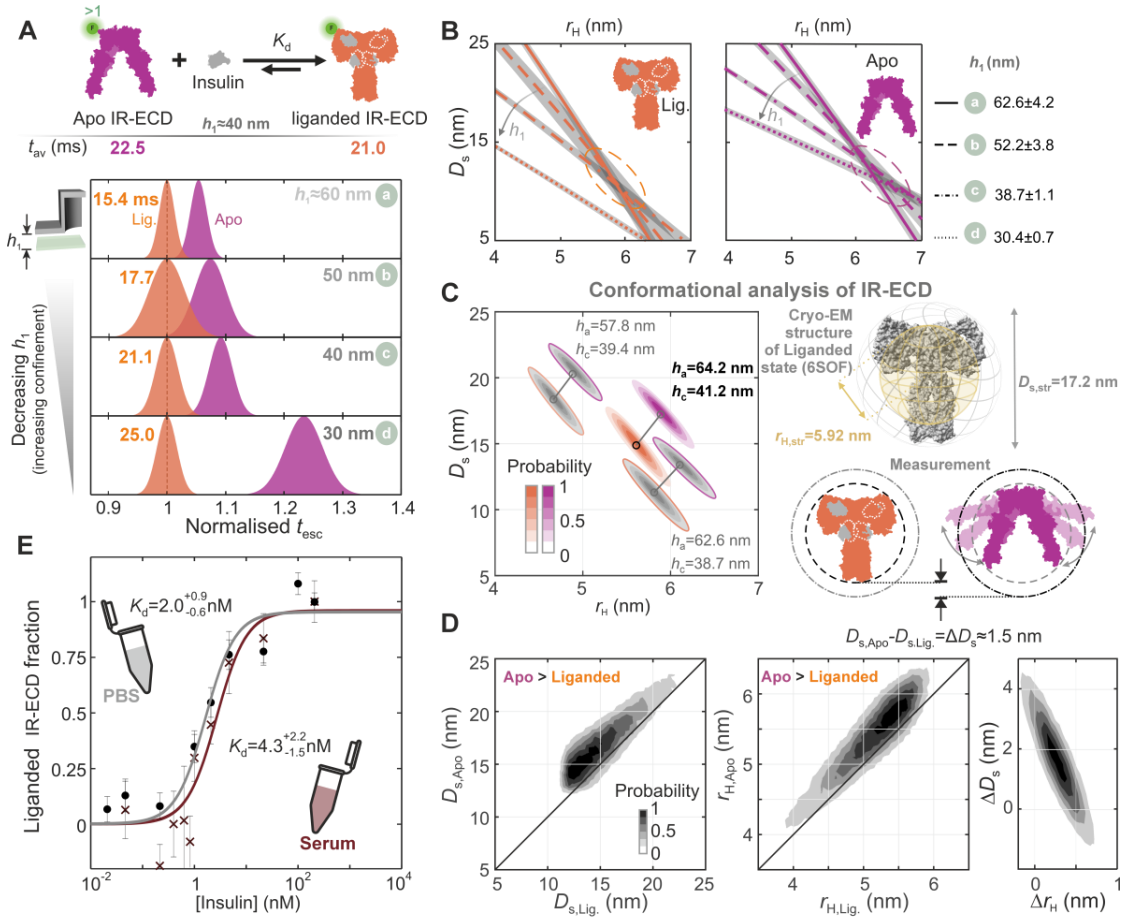


Fig. 7. Detection and structural modelling of conformation changes of the Insulin Receptor Ectodomain (IR-ECD) upon insulin binding. (**A**) Upon incubation with 100 nM insulin, well above the expected K_d of the interaction reported in Fig. 6B, the fluorescently labelled IR-ECD displays a reduction in average escape time t_{av} from 22.5 ± 0.3 ms to 21 ± 0.1 ms in a device where h_1 ≈40 nm, used for K_d measurements in (**E**). Examining the ligand-free 'apo' and the liganded states of the IR-ECD in slits of height ranging from $h_1 \approx 60$ to 30 nm reveals progressively increasing fitted t_{esc} values for both states, as well as a systematically increasing ratio of the two timescales, clearly indicating significant differences in 3D conformation between the two states (plots a-d). (**B**) Graphical presentation of Eq. (1) for the liganded state (left) and 'apo' state (right) using fitted t_{esc} values in slits with quoted h_1 values determined by calibration (see fig. S13). Lines represent (D_s, r_H) solutions expected for $h_1$corresponding to the mean value obtained from. Shaded grey bands denote uncertainties given by the precision of the t_{esc} measurement. Similar to Fig. 3, the intersection of any two bands gives the range of allowed values of (D_s, r_H) characterizing the molecular state. (**C**) Three possible sets of solutions describing the 'apo' and liganded states obtained by assuming h_1 values that vary within one standard deviation of the mean values measured for devices 'a' and 'c' (left). Note that the true uncertainty in slit height (<ca. 1 nm) is expected to be much smaller than that obtained by the calibration procedure (*20*). Although absolute values for D_s and r_H are not determined with high accuracy at present, the difference between the two states is systematic and points to an increase of $\Delta D_s \approx 1.5$ nm and

$\Delta r_{\mathrm{H}} \approx 0.25$ nm for the 'apo' state compared to liganded case. Conceptual view of the inferred conformational difference between the 'apo' state and the cryo-EM ligand saturated state (6SOF) as inferred from ETsy measurements (right). (**D**) Plots of $D_{\mathrm{s,Apo}}$ vs. $D_{\mathrm{s,Lig.}}$ and of $r_{\mathrm{H,Apo}}$ vs. $r_{\mathrm{H,Lig.}}$ for several different pairs of h_1 values lying within the range inferred from calibration for the (a, c) device geometry pair. The plots confirm a systematic compaction in overall extent of the liganded state compared to the ligand-free 'apo' state (fig. S13). (**E**) Affinity measurement for the binding of the IR-ECD to insulin. The bound fraction is inferred from the reduction in average timescale t_{av} of labelled IR-ECD due to the conformation change upon insulin-binding, as shown in (**A**) (see Fig. 6, fig. S19 and supplementary text for details). Measurements in both PBS (black data) and in simulated human serum (red data) give similar K_{d} values that are consistent with the literature.

Statement of Authorship for joint/multi-authored papers for PGR thesis

To appear at the end of each thesis chapter submitted as an article/paper

The statement shall describe the candidate's and co-authors' independent research contributions in the thesis publications. For each publication there should exist a complete statement that is to be filled out and signed by the candidate and supervisor **(only required where there isn't already a statement of contribution within the paper itself).**

Title of Paper	Single-molecule size-and-shape spectrometry on a chip
Publication Status	□ Published □ Accepted for Publication x Submitted for Publication □ Unpublished and unsubmitted work written in a manuscript style
Publication Details	Xin Zhu†, Timothy J. D. Bennett†, Konstantin C. Zouboulis, Dimitrios Soulias, Michal Gryzbek, Justin L. P. Benesch, Afaf H. El-Sagheer, Uenal Coskun and Madhavi Krishnan †authors contributed equally

Student Confirmation

Student Name:	Timothy Bennett		
Contribution to the Paper	Contributed to development of methodology. Performed experiments and analysed data. Contributed to the preparation of the manuscript and wrote the supporting information.		
Signature	TBennett	Date	09-10-2024

Supervisor Confirmation

By signing the Statement of Authorship, you are certifying that the candidate made a substantial contribution to the publication, and that the description described above is accurate.

Supervisor name and title: Professor Madhavi Krishnan			
Supervisor comments			
Signature		Date	10/10/2024

This completed form should be included in the thesis, at the end of the relevant chapter.

Chapter 5

Conclusion

In this thesis a series of improvements and applications of ET*e* have been explored. The escape-time technology has been taken from ensemble measurements of effective charge under low salt conditions, to single-molecule measurements of effective charge under near-physiological, and size and shape under physiological conditions. The implications of such improvements are that escape-time chips now provide a versatile toolbox for a broad range of biological applications.

Real-time intermolecular interaction monitoring

The ability to distinguish size heterogeneity in a sample, combined with short measurement times, enables the monitoring of reaction progress and measurements of binding affinities and kinetic constants. The low sample volumes and concentrations (picolitres of picomolar samples), and lack of surface tethering provide considerable advantages over laboratory standard techniques such as surface plasmon resonance (SPR), isothermal titration calorimetry (ITC), biolayer interferometry (BLI) or microscale thermophoresis (MST)[103]. Note that the sample concentrations referred to here are that of labelled material which can be orders of magnitude below that of the unlabelled sample, enabling quantification of intermolecular interactions up to to millimolar in affinity.

Quantification of conformational heterogeneity and oligomerisation state

The ability to distinguish size, charge and shape heterogeneity with single-molecule precision enables quantification of the equilibrium distribution of states in a complex sample. Such measurements can provide differentiation of conformation states in a sample, which could be applied to the variety of stable tertiary structures common in single-stranded RNA under native conditions. Combined with the short measurement times, the kinetics of stochastic state-switching in a dynamic equilibrium of these conformations could also be probed. Probing these dynamics

typically requires single-molecule methods such as FRET[104,105], however with just a single non-specifically positioned fluorophore, and no surface tethering required, an escape-time measurement has its advantages.

Proteomics, genomics and transcriptomics on a chip

The ability to localise fluorescent signal from single molecules while interacting with the fluid traps could enable sample fingerprinting through multiplexing. Spectrally distinct fluorescent probes, which are highly specific to a target region of a sample, facilitate spectral multiplexing. By co-localising fluorescence signal from just a few distinguishable probes bound to the target species using multi-channel fluorescence imaging, a multitude of unique species can be identified. An escape-time realisation of this could also enable another dimension of multiplexing with the use of differently-sized probes, which would give variability in escape-time signatures as well as spectral distinguishability. The multiplexing potential combined with the single-molecule resolution offers quantitative characterisation of a sample, which is advantageous over conventional methods such as microarrays, fluorescence in-situ hybridisation (FISH), or quantitative polymerase chain reaction (qPCR) methods where additional biochemical reactions are often required for signal amplification[106]. The single reaction step of an escape-time method would also give considerable advantages over other single-molecule methods, many of which rely on sandwich assays in multi-step, laborious protocols, such as single-molecule array (SiMoA) assays[107].

An additional dimension which ET*e* offers is the ability to detect small charge modifications, which could enable epigenomic and PTM-omic assays when chemical modifications are charge altering. Glycosylation for example, is a PTM which is linked to a number of diseases - the levels of glycated haemoglobin can be a biomarker for both type 1 and type 2 diabetes for example[108]. Such a modification induces both a modest size shift, and more importantly a charge shift, and thus glycan profiling is well suited to an escape-time measurement.

Rare-species detection for diagnostics

The high-throughput quantitative molecule counting could enable detection of rare-species within a heterogeneous sample when the target species can be differentiated. Specific covalent labelling with a fluorophore (such as using the warhead-linked probes used for activity/affinity-based proteomics[109], or the oligonucleotide-linked probes used for genomics and transcriptomics), enables such detections from the change in size of the 'bait' probe upon binding to the target species. Because ET*e* can offer order of magnitude shifts in timescale upon binding of the

probe to the target, the unbound probe can be readily screened out unlike other methods - this embodies the power of a measurement rather than just a detection technology. We envisage such strategies being used in a diagnostic assay for the detection of biomarkers down to femtomolar concentrations (currently, but with further throughput enhancements, attomolar concentrations), against a complex background such as blood serum. Such measurements have direct applications to the detection of disease-states, many of which are characterised by protein or nucleic acid biomarkers present in human blood at low nanomolar to attomolar concentrations[110,111]. The small sample volumes and low detection limits rival conventional enzyme-linked immunosorbant assay (ELISA) or qPCR methods, without the need for complex and laborious protocols or sample amplification[106,110].

Detection of binding-induced conformation changes for pharmaceuticals

The sensitivity to different conformations allows detection and measurement of binding-induced conformation changes. Escape-time measurements offer coarse-grained detection of relative differences in structural properties, which is an under-exploited dimension in intermolecular interaction screening, since the binding of small ligands which negligibly increase mass, can induce significant conformation changes. Such measurements could be useful for the detection of riboswitch switching for example, or the release of iron from transferrin, both of which change conformation measurably upon the addition/departure of a small molecule or atom with a fraction of the mass of the macromolecule[17,112,113].

Additionally, since conformation is so inextricably tied to function in nature, such measurements could be applied to drug-screening, where induced conformation change can be used as an indicator of not only binding, but also of inhibited function. This additional modality offers advantages over other techniques often employed for drug-screening which rely upon solely the detection of binding, such as affinity selection mass spectrometry (AS-MS), thermal shift assays (TSA), SPR and MST[114].

Structural biology on a chip

The measurements of charge and shape enable inferences on absolute structural dimensions. With the addition of modelling, escape-time measurements can inform fine-grained estimates of structure under native conditions. Such measurements in the solution phase, at room temperature, and without the need for stabilising ligands, surface tethering or complex analysis offer advantages over laboratory standard methods such as X-ray crystallography, cryo-EM or small-angle x-

ray scattering (SAXS). While acknowledging the relatively limited precision and the need for additional calibration to account for various experimental inaccuracies, the ease of escape-time measurements make them an appealing alternative to such infrastructure-expensive techniques.

Overall, ETe and ETsy offer a versatile range of applications from measurements of effective charge, size and shape, facilitated by fast measurements, high precision and single molecule tracking, all done in the solution phase under native conditions. We can learn a lot from simply watching molecules diffuse!

Appendix A

Device fabrication

Three types of devices were used in this study: Gen 1 ETe devices (used in experiments in Chapter 2 and section 3.1), Gen 2 ETe devices (used in section 4) and FlETe devices (used in section 3.2). The fabrication procedure for original ETe devices has been previously reported[93], and uses a combination of photo- and electron-beam lithography to pattern the devices on a silicon wafer, followed by reaction ion etching (RIE) to define the nanostructure. Once the silicon part has been realised, SEM and AFM scans of the device reveal the pocket and channel dimensions, before anodic bonding is used to seal the channels with a glass coverslip.

The design of FlETe devices is similar to Gen 1 ETe devices with the replacement of pockets with line-traps, as shown in Fig. 17. A re-useable seal made out of polydimethylsiloxane (PDMS) was used to integrate the device into a microfluidic circuit to apply the vacuum.

The design of Gen 2 ETe devices was changed to facilitate 4-inch silicon wafer-scale thermal bonding of multiple devices simultaneously. The thermal bonding enabled shallower channel heights to be obtained ($2h < 60$ nm), and the yield was greatly improved with wafer-scale bonding. The fabrication process for these devices uses laser-lithography for patterning, and a combination of RIE and inductively coupled plasma RIE (ICP RIE) to define the nanostructure Fig. 18. Devices were directly integrated into a microfluidic circuit to apply the vacuum.

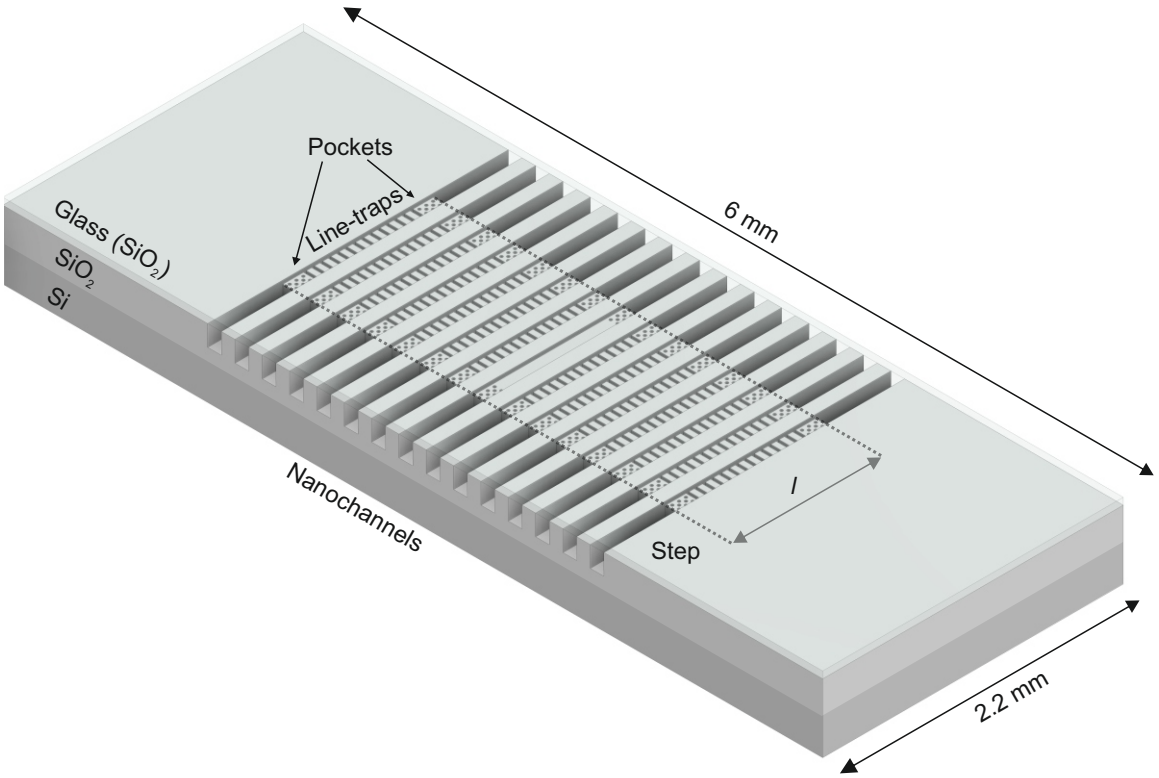


Fig. 17 **Schematic of a device used for FlET*e* experiments.** A multitude of 5 μm wide nanochannels run along the the width of the device. Deeper "step" regions at either end of the channels (ca. 400 nm deep), constrain the shallower "slit" part of the device (ca. 70 nm deep) to a length l. Pockets ca. 300 nm deep and ca. 500 nm in diameter are patterned at either end of the slits to coincide with observation windows. Line-traps ca. 300 nm deep and ca. 500 nm wide are patterned in the rest of the slit region, spaced with edge-to-edge distance $L \approx 1.5\,\mu$m. Diagram not to scale.

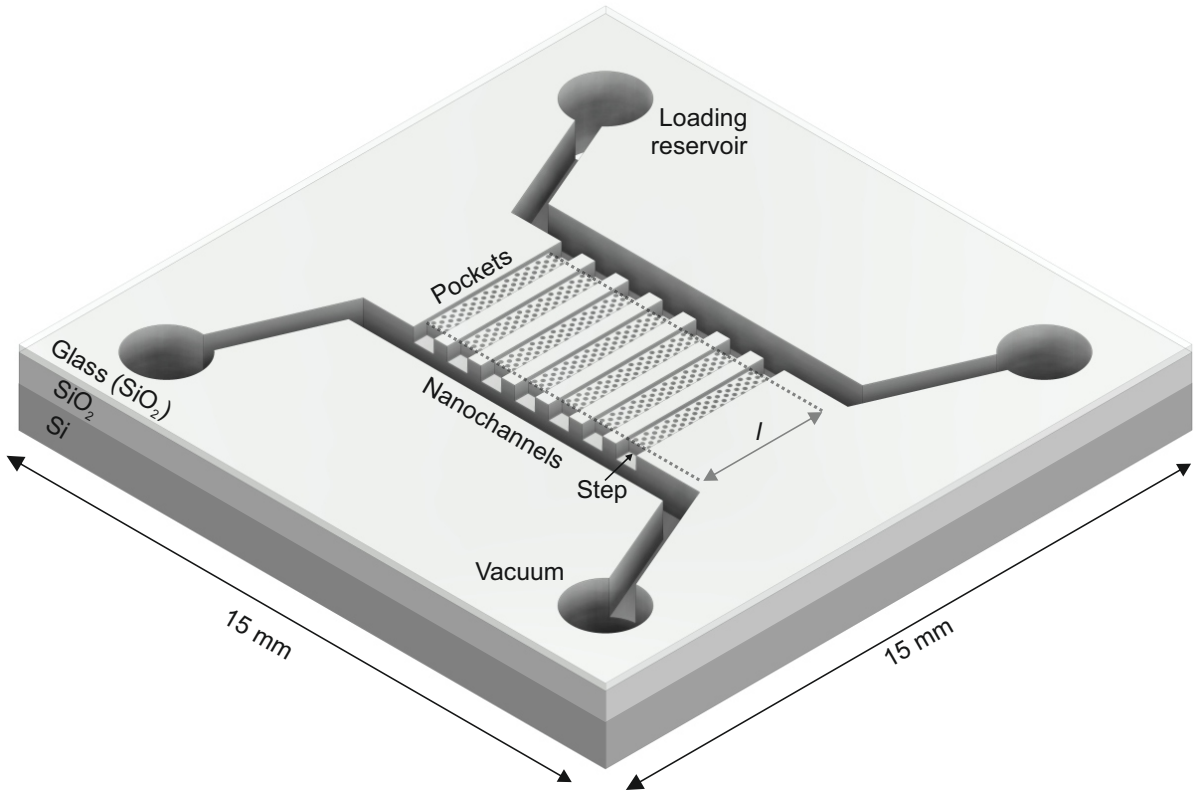


Fig. 18 **Schematic of a device used for FlET*e* experiments.** A multitude of 5 μm wide nanochannels run between two deep feeder channels loaded with sample on one side, and connected to vacuum on the other. Deeper "step" regions at either end of the channels (ca. 400 nm deep), constrain the shallower "slit" part of the device (ca. 70 nm deep) to a length *l*. Pockets ca. 300 nm deep and ca. 550 nm in diameter are patterned throughout the slit regions. Diagram not to scale.

Appendix B

Widefield epifluorescence microscopy

Three different microscope setups were used for the studies shown in this thesis. The measurements in sections 2.1 and 3.1 used an "original" ET*e* setup as outlined in Fig. 19, while measurements in sections 2.2, 3.2 and 4.1 used the "upgraded" ET*e* setup in Fig. 20. A third setup specifically for FlET*e* was also built, see Fig. 23.

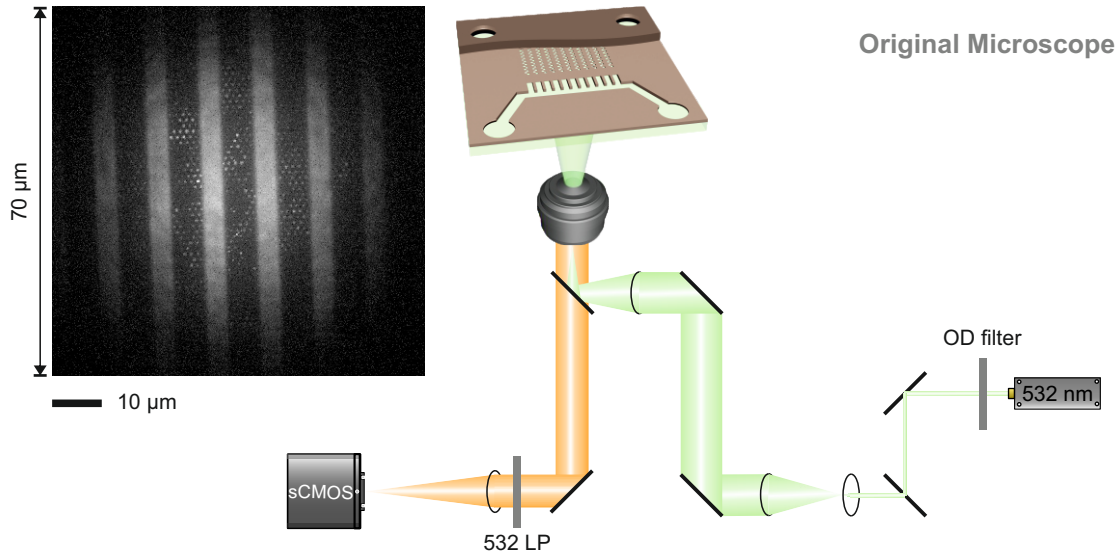


Fig. 19 **Diagram of the original microscope.** A standard epi-fluorescent widefield fluorescence microsope for excitation at 532 nm, and imaging a FOV of $\lesssim$ 100 µm at $\lesssim$67 Hz. A maximum intensity image of a typical measurement of 60 bp DNA is shown.

The precision of an ensemble ET*e* measurement is dictated by the total number of escape events measured, N. To increase N for a given measurement time, more events need to be measured simultaneously. This is what the upgraded ET*e* setup aimed to address: a larger microscope field-of-view (FOV) was required while maintaining the signal-to-noise ratio (SNR). The key improvements in the setup are outlined here.

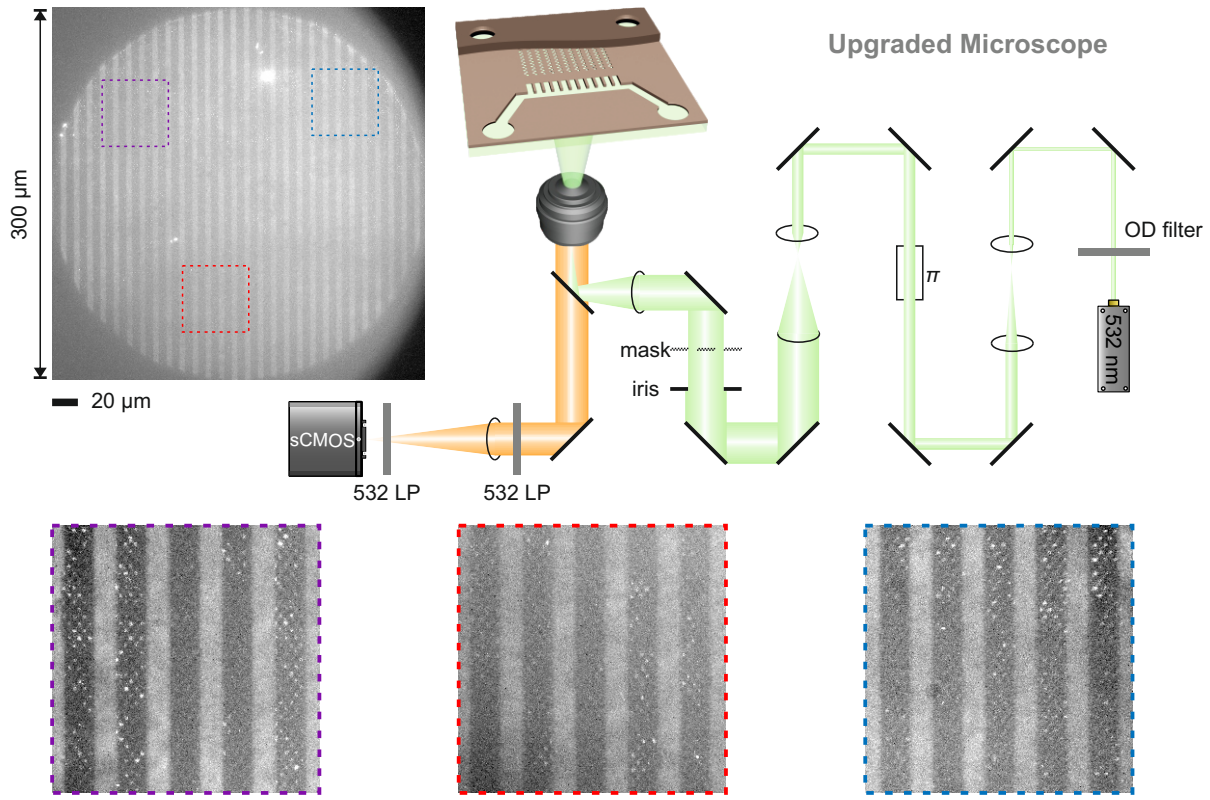


Fig. 20 **Diagram of the upgraded microscope.** Beamshaping with a pi-shaper (π) of a higher powered laser combined with optimised objective alignment enabled illumination of a large FOV. A water objective, cropping of the excitation pathway and a high speed camera enable high SNR measurements over $\lesssim 330$ µm at $\lesssim$100Hz. An additional mask can be used to isolate observation windows used in FlET*e* measurements. A maximum intensity image of a typical measurement of 60 bp DNA is shown, along with zoomed regions illustrating the uniformity of the FOV.

Excitation pathway

A higher power 3 W laser was obtained (DMPV-532-3, Del Mar Photonics), and to further improve the usable area of the FOV a beam shaper was employed to flatten the intensity profile of the laser. This was crucial for single-molecule tracking experiments as a low power density is required to maximise the number of escape events measured before photo-bleaching - the power density is better controlled with a uniform, flat illumination profile. The beam shaper used was a πShaper 6_6 (AdlOptical), which turns a ca. 6 mm Gaussian profile into a top hat profile. To illuminate the largest FOV possible on the device, either high angles of incidence are required for the focused excitation beam on the back focal plane of the objective, or a wide incident beam is required. The physical constraints of requiring a dichroic mirror before the objective limited the focal length of tube lens possible, hence the former was hard to achieve. As a result, a wide incident beam-width approaching 2 inches was required. To avoid burning of the objective, the incident beam was deliberately de-focussed to approximately 1 cm below the objective casing - as a result the illumination profile was divergent out of the objective.

Background reduction

A water immersion objective was used (60x, 1.2 NA, Olympus), this was found to increase the SNR of our measurements compared to an oil counterpart with higher numerical aperture (NA), see Fig. 21. We attribute this to the rejection of background signal from emitters in the glass which would only be detected with higher NA objectives. This observation is contrary to many single molecule experiments where the fluorophores of interest are deposited on glass surfaces, in that instance oil immersion and higher NA is usually the preferable.

We also observed that the background noise increased with the total illuminated area of the sample, while the fluorescent signal remained as a function of power density only. We rationalise this observation by considering the main source of background as autofluorescence of the silica device whose sources are evenly distributed throughout the device. The light cone with opening angle defined by the NA describes the sources of photons collected by a point on the detector, see Fig. 22. Light emitted by a point source fluorophore at the focal plane is collected to this point on the detector, but also contributions from out of focus emitters which are illuminated in the cone behind the point source. The total area illuminated determines the extent of these additional contributions, without altering the signal detected by the in focus emitter, thus background noise is independent of fluorophore signal. To combat this, a compromise was empirically determined

by using an iris in the excitation pathway to crop the beam, achieving the largest possible illumination area while maintaining good SNR. This turned out to be ca. 300x300 μm, compared to the maximum possible of ca. 420x420 μm.

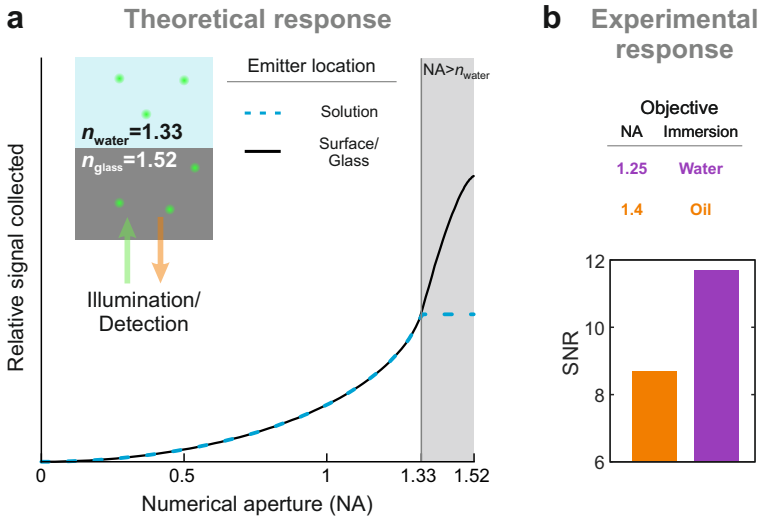


Fig. 21 **Justification for lower NA, water immersion objectives giving higher SNRs.** (**a**) Illustrates the theoretic response from emitters located in solution (e.g. fluorescent molecules, blue dashed line), vs located in the glass coverslip (e.g. colour centre/autofluorescence, black line) collected using different numerical aperture (NA) objectives. When NA$>n_{\text{water}}$ no additional response is expected from emitters in the solution (molecule signal) compared to emitters in the glass (background). Theoretical calculations are from Mahanta et al.[115]. (**b**) A measurement of SNR from an ET*e* measurement of 60 bp DNA using two different 40x objectives: 1.25 NA water immersion (40x Nikon, purple data), 1.4 NA oil immersion (40x Zeiss, orange data).

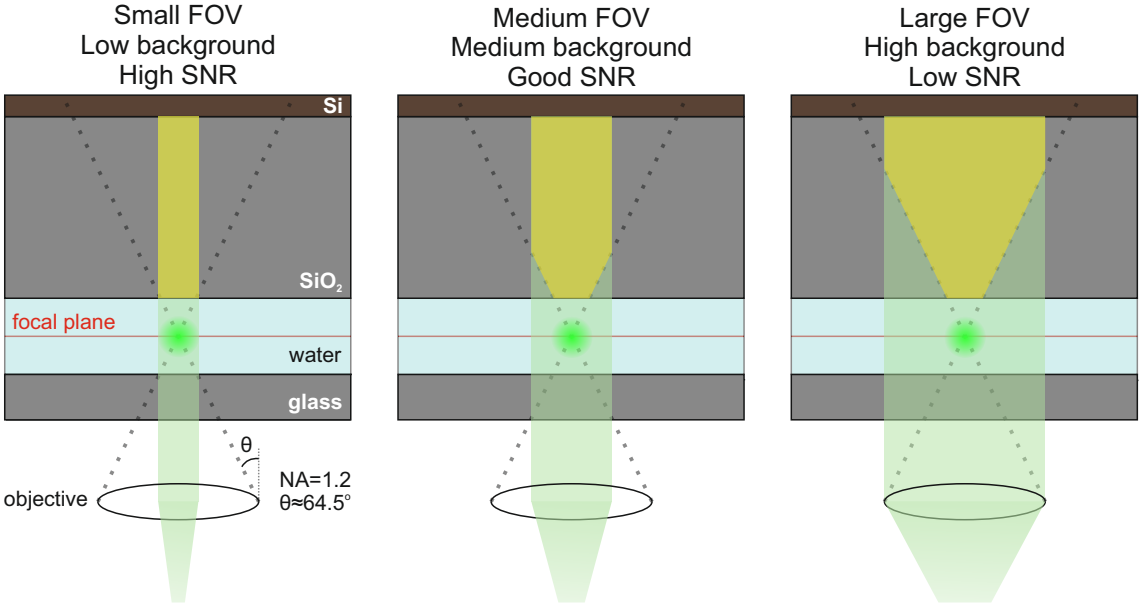


Fig. 22 **Sources of background in an ET*e* device and their dependence on illumination area.** Illuminated area (shown by green rays) produces a constant signal from an in-focus emitter on the focal plane (red line) in the solution, regardless of illuminated area. As the illuminated area increases (left to right), the number of background emitters in the SiO_2 device increases (yellow area).

Emission pathway

A high pixel number, high speed sCOMS camera was used for imaging (Kinetix, Photometrics). This allowed reading out of 1400 lines (210 μm on the device) in 5 ms, as required by the majority of experiments. A restriction on the magnification of our experiments is the requirement to project the image of each pocket onto at least 1 pixel (ideally ca. 3x3 grid of pixels), to ensure uniformity in signal across all pockets. In the limiting case of equality between pocket diameter and pixel size, then in some instances the photons from a trapped molecule could be spread over a 2x2 grid of pixels - in this case a 4 times lower SNR would be observed than if the pocket perfectly lined up with a pixel. In the setups in Fig. 19 and Fig. 23 the image pixel size was ca. 120 nm, and in the setup in Fig. 20 was ca. 150 nm. In principle a single pixel per pocket is all that would be required for imaging, given perfect alignment between pixels and pockets, which could massively decrease the specification of a camera required.

Microscopes for FlET*e*

FlET*e* measurements require two relatively small observation windows separated by a relatively large distance. The upgraded microscope was used for these measurements due to the large FOV, and the space between the two observation windows was obscured from illumination using a custom mask, see Fig. 20. This method allowed for window separation limited by the total imageable field of the view of the microscope. Since observation of the entire field in between windows is not required, an alternative microscope was designed which allowed simultaneous imaging of two arbitrarily spaced regions of interest (ROIs), see Fig. 23. The excitation was split by a polarising beamsplitter and a small lateral shift was induced in one beam, such that once recombined with the unaltered beam, it passed through the objective at a slightly tilted angle, illuminating a different spot on the device. This meant that a lower power 200 mW laser (Photontec, Berlin), and one-inch optics could be used while still illuminating the observation windows with the same power density. This offers a cheaper microscope design to the upgraded setup. The imaging pathway was the same and so a fundamental limitation of the separation of the two observation windows still remained from the field number of the objective.

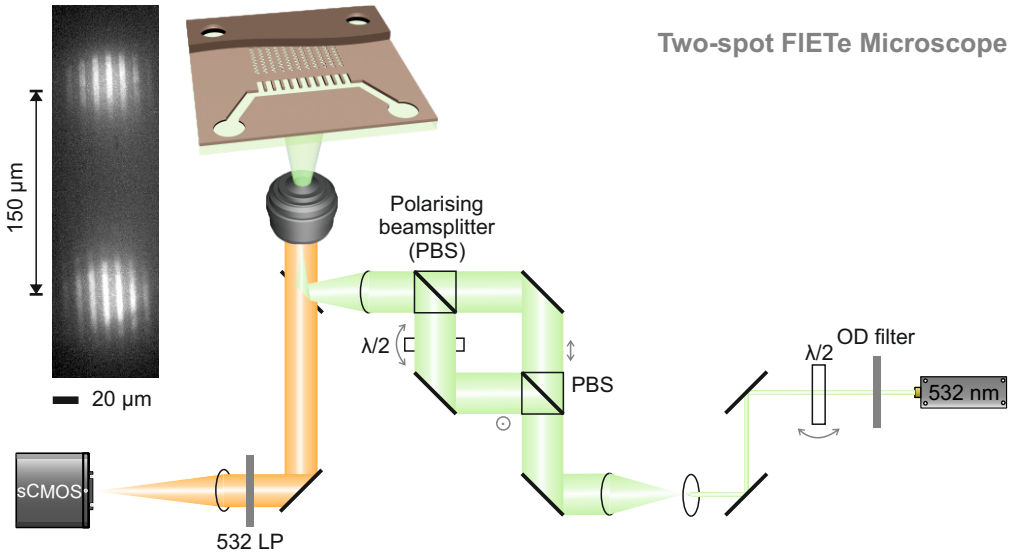


Fig. 23 **Diagram of the two-spot FlET*e* microscope** A half wave plate (λ/2) is used to rotate the polarisation direction of the laser to be at 45°to a polarising beamsplitter. The beamsplitter is used to add a small lateral shift in the position of 50% of the beam, which illuminates a separate ROI on the device. These two ROIs are aligned with observation windows in a FlET*e* device for a measurement. A maximum intensity image of a typical measurement of 60 bp DNA is shown.

Appendix C

Escape-time movie analysis

Escape times, t_{esc}, are extracted from raw movies using custom MATLAB scripts. Briefly, locations of pockets are detected in a given movie from the maximum intensity projection; because the molecule diffuses rapidly when not confined to a pocket (>1 μm in a 5 ms exposure time), it is essentially invisible to the camera, thus any pixels with high signal correspond to positions where a molecule has been localised for >5 ms - the traps. Intensity-time traces of each pocket are analysed to extract the individual escape events, Δt. A histogram of these times is created and fit with Eq. 12.

C.0.1 Detection of pockets

Pocket locations are generally detected from the maximum intensity projection using the particle detection software TrackNTrace[94]. However in some measurements pockets detected by TrackNTrace were spurious bright pixels in the FOV which did not correspond to molecules trapped in pockets. These could be caused by non-specific binding of the molecules to the channel surface, "sticking", or colour centres in the silica among other things. In these cases two different methods were employed, outlined below.

Classification of pockets

Due to the regular lattice of pockets in the channels, it was possible to identify spurious detections in off-lattice locations by post-processing. Each detected pocket was categorised based on the number of adjacent detections which were correctly spaced using a nearest neighbour algorithm. For example, if all pockets were correctly detected, each detection would have six nearest neighbour detections at the same distance apart, since the pockets are arranged in a hexagonal lattice. A threshold of two correct detections was usually applied. This method relies upon the majority of

detections being correct, "on-lattice", and also relies on a high density of pockets with molecules trapped in.

Importing of pocket locations

Some molecules exhibited a lot of sticking, resulting in lots of off-lattice pockets detected. When lots of detections are incorrect, the previous nearest-neighbour method breaks down. In these instances, prior to measurement of the sample of interest (A), the "sticky" molecule, a sample of a "non-sticky" molecule (B), could be measured in an identical FOV. The pocket array detected from the measurement of sample B could then be imported into the analysis of sample A, and only escape events from pockets detected for both A and B considered. A small amount of movement of the FOV between consecutive measurements was accounted for by manually aligning the imported pocket array with visually on-lattice regions of the measured pocket array for sample B.

Similarly, where experiments required low sample concentration, e.g. TrETe measurements, the nearest-neighbour method excluded many correct pockets due to a lack of neighbouring detections. In this instance a high concentration measurement done prior on an identical FOV could be used to identify pocket locations, which could then be imported in the low concentration measurements.

C.0.2 Measurement of Δt

Signal from arrays of pixels corresponding to each pocket are pooled, and two thresholds are set to discriminate a signal emitted from a molecule confined in the pocket versus background noise. The area of pixels pooled for each pocket roughly aligns with the escape radius criterion applied to the BD simulated particle trajectories, discussed in Appendix D. Colloquially, the thresholds identify frames likely to contain a trapped molecule, and frames which definitely contain a trapped molecule. Escape event length is dictated by the former, but the event is only exported to the t_{esc} histogram if at least one frame of the event satisfies the latter. Examples of this intensity-time trace analysis is shown in Fig. 26**b** in Appendix F.

Background determination

Thresholds are set based on a number of standard deviations of the noise above the background level, thus it is crucial to quantify the background signal accurately. This is done independently for each pocket; the pooled pixel signal is combined on a rolling basis of ca. 500 frames, a histogram

is constructed and fitted with a sum of two Gaussians - one corresponds to the background and the other corresponds to the signal peak. The mean and width of the smaller intensity Gaussian is taken to be the background and noise for that pocket during that time window. This was particularly important for long measurements, such as TrETe and FlETe, where background was sometimes seen to vary due to laser fluctuations or external factors. The SNRs reported variously throughout correspond to the signal relative to these locally determined background and noise levels.

Appendix D

Brownian dynamics (BD) simulations of escape-time experiments

D.1 BD simulations of ET*e* experiments

As outlined in section 1.4.3, BD simulations are used to find the exact relation between an experimentally measured t_{esc} and the energy well depth of the trap, W, for a given molecule diffusion coefficient D_{t}. Details of the simulations are given here.

D.1.1 Simulation of particle trajectories

A 2D simulation of the trapping environment is setup: an x-y plane confined at x=0 and x=5 μm (for a typical channel width) with no explicit pocket geometry imposed, and radially symmetric potential wells imposed, centred on the pocket locations. The potential wells have energy depth W, which is varied to track the residence times of particles diffusing with diffusion coefficient D_{t}. The 2D simulation was confirmed to fully reproduce the results obtained from a 3D simulation in previous work[75], while requiring far reduced computation time.

W in an ET*e* experiment consists of two factors, as expressed in Eq. 13: ΔF_{el} and f. The former is a linear function of ϕ_{m}, while the latter comes from the implicit fluctuation of the particle from the midplane of the device, thus the shape of ϕ along the midplane of the device is used for the potential well in the 2D simulation. $\phi(r, z)$ is calculated by solving the PB equation (Eq. 6) using COMSOL multiphysics for our cylindrically symmetric device geometry, and the minimum $\phi(\mathrm{r}, \mathrm{z})$ at each r value yields the shape along the midplane, $\phi_{\mathrm{mid}}(r)$, as shown in Fig. 2. A 360° rotation of $\phi_{\mathrm{mid}}(r)$, yields $\phi_{\mathrm{mid}}(x, y)$, which is centred on the pocket coordinate in the x-y plane and rescaled to be zero at the centre of the pocket, and W outside: $\phi_{\mathrm{mid}}'(x, y)$.

A particle trajectory through the landscape is generated by iteratively solving the Langevin equation for light particles, expressed in 1D as Eq. 23.

$$x(t+\delta t) = x(t) - \frac{D_{\mathrm{t}}\,\delta t}{k_{\mathrm{B}}\,T}\,\nabla\phi_{\mathrm{mid}}{}'(x) + \sqrt{2\,D_{\mathrm{t}}\,\delta t}\,\mathrm{F}(t) \tag{23}$$

$\mathrm{F}(t)$ represents the stochastic motion of the particle, and is drawn randomly from a normal distribution with mean 0 and standard deviation 1, such that $\langle \mathrm{F}(t)\rangle = 0$ and $\langle \mathrm{F}^2(t)\rangle = 1$. δt is the simulation timestep, $\delta t = 0.1\,\mu\mathrm{s}$ was used for all simulations in this thesis.

The particle trajectories are sampled to match the experimental conditions and analysed using two escape criteria to extract the lengths of residence times, Δt, inside the traps before escape. The escape criteria are determined empirically through comparison with synthetic movies of the same trajectories, outlined in more details in the following section D.1.2. These residence times are histogrammed similarly to the experimental movies to extract t_{esc} (see section 1.4.3), and thus the relation between t_{esc} and W can be determined for a given D_{t}.

D.1.2 Synthesis of movies of particle trajectories

The particle trajectories are made into movies, mimicking those from experiments, in order to account for imaging artifacts in extracting Δt experimentally. Particle trajectories are sampled in line with experimental conditions by evaluating the mean and standard deviation in particle position over all the simulation frames equating to an experimental exposure frame (typically 5 ms in length). The molecule signal in each synthetic movie frame is given by a 2D gaussian with x and y widths given by the product of the standard deviation in particle position and microscope point spread function (PSF). The amplitude of the signal is chosen from a Gaussian distribution matching the characteristics of the experimental movies, exemplified in Fig. 26**a** in Appendix F. This signal is centred on the particle position in each frame, and added to a background. The background has two manifestations of noise - spatial and temporal, which both combine to give that characteristic of an experiment. Such movies are analysed similarly to experimental movies to extract Δt and t_{esc}.

Making these movies is computationally expensive, and so empirical criteria are determined to enable analysis of the particle trajectories directly. Two radial escape thresholds are used, $R_{\mathrm{esc},1}$ and $R_{\mathrm{esc},2}$, and similarly to the two intensity thresholds used for event detection in a movie (see section C.0.2), they identify frames likely to contain a trapped molecule, and frames which definitely contain a trapped molecule. Escape event length is dictated by particle position being

inside the former, but the event is only exported to the t_{esc} histogram if at least one frame of the event is inside the latter, see Fig. 24. The choice of value for these escape criteria is such that t_{esc} as determined from the trajectory using these criteria is indistinguishable from the t_{esc} as determined from the synthetic movie of the same trajectory. Escape thresholds of 200 nm and 350 nm were used for the BD simulations informing experiments in Chapter 2.

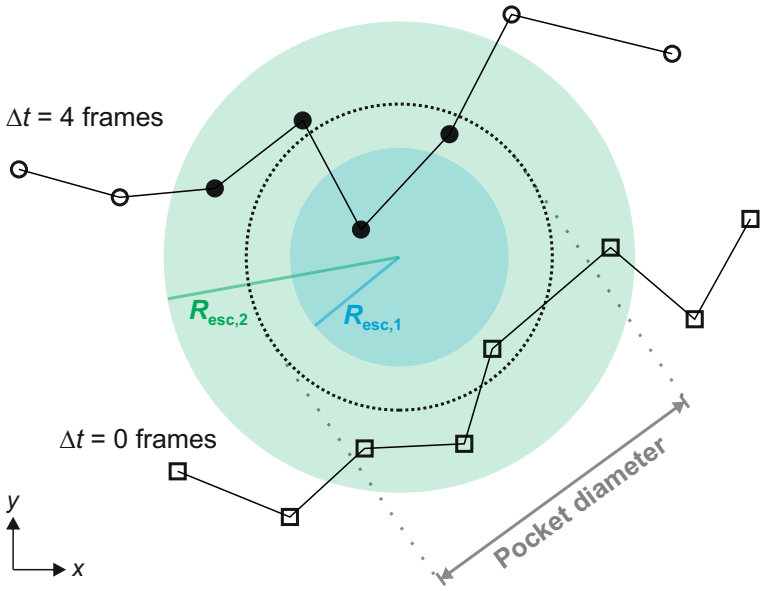


Fig. 24 **Schematic of escape criteria applied to 2D BD simulations.** Two hypothetical 2D particle trajectories are shown (circles and squares) in the vicinity of a pocket (black dashed line). The trajectory indicated by circles, ventures inside $R_{esc,1}$ for at least 1 frame and so an escape event is recorded of duration given by the time the particle spent inside $R_{esc,2}$ (4 frames). The trajectory indicated by squares, does not venture inside $R_{esc,2}$ and so no escape event is recorded. despite there being 4 frames when the particle was inside $R_{esc,2}$.

D.2 BD simulations of FlET*e* experiments

To better understand FlET*e* experiments and to explore the impact different parameters has on our experiment, a series of Brownian Dynamics (BD) simulations were setup of the experiment. Initially 3-dimensional simulations were setup of a point particle diffusing, with diffusion coefficient D_t, along a 5 μm wide channel with regularly spaced rectangular line-traps. A flow field for a single trap and surrounding channel was calculated using the microfluidics module in COMSOL and superimposed on each trap. Similarly an electrostatic potential was calculated using COMSOL and superimposed on each trap. The time taken for a particle to migrate a set distance down a channel was recorded, t_{FlETe}.

These "full-monty" simulations were very computationally costly and so to streamline the process, a 1-dimensional approximation of the system was constructed. Assuming the molecule

sits on the midplane of the channel and pocket (and far enough away from the channel walls) the midplane electrostatic potential and centerline flow velocity can be used to simulate the progress down a 1D channel. We note here that similarly to the 2D BD simulations used in the previous section (D.1) the midplane electrostatic potential must be rescaled to include the contributions of fluctuations to the total energy well depth, f in Eq. 13, since we have no explicit z oscillation of the particle. f is calculated as outlined in section 1.4.3. Fig. 25 shows the comparison of these two simulations for particles with diffusion coefficient $D_t = 5x10^{-7}$ cm^2/s (as for 60 bp DNA). We see that the 1D simulation matches the 3D simulation for $\Delta F_{el} \gtrsim 1$ k_BT, corresponding to $W \gtrsim 3k_B$T; below this value the particle is regularly sampling space closer to the channel surfaces and as a result is experiencing slower flow rates, hence the slower progress down the channel in the 3D case than 1D - t_{conv} is increased. We note that at the limit of the no electrostatic potential the 3D simulation matches the 1D simulation using the average flow speed rather than the centreline velocity, since here there is no bias of the particle towards the centre of the channel so the convection speed is set by the average fluid velocity. Note that with no electrostatics the $W = 1.35\, k_B\, T$ due to the entropic well depth, see Fig. 4.

$W \gtrsim 3\, k_B\, T$ corresponds to a molecule with $q_{eff} \lesssim -10\, e$ in a typical trap geometry ($2h = 70$ nm, $d = 200$ nm with $\phi_s = 2.7\, k_B\, T/e$ in $c = 1$ mM NaCl) and so is valid for all experiments presented later, thus 1D simulations were employed to study the effect of Péclet number, as outlined in section 3.2.2.

D.3 BD simulations of ETsy experiments

To verify the theoretical response of t_{esc} to D_t and D for a given h_1 and h_2, given by Eq. 22, a series of BD simulations of spherical particles in different trap landscapes were setup. Contrary to in sections D.1 and D.2 where lower dimension simulations could be setup, it is exactly the additional the fluctuation free energy term ΔF_{trans} from Eq. 13 that we are interested in probing here, thus we require a 3D simulation.

3D geometries were constructed of a single trap, and the trajectory of a spherical particle, radius R, was simulated by iteratively solving the Langevin equation for light particles in the absence of any potentials, $\phi_{mid}'(x, y, z) = 0$ in Eq. 23. The particle's diffusion coefficient was set according to its radius using the Stokes-Einstein-Sutherland equation, Eq. 24.

$$D_t = \frac{k_B\, T}{6\, \pi\, \eta\, R} \tag{24}$$

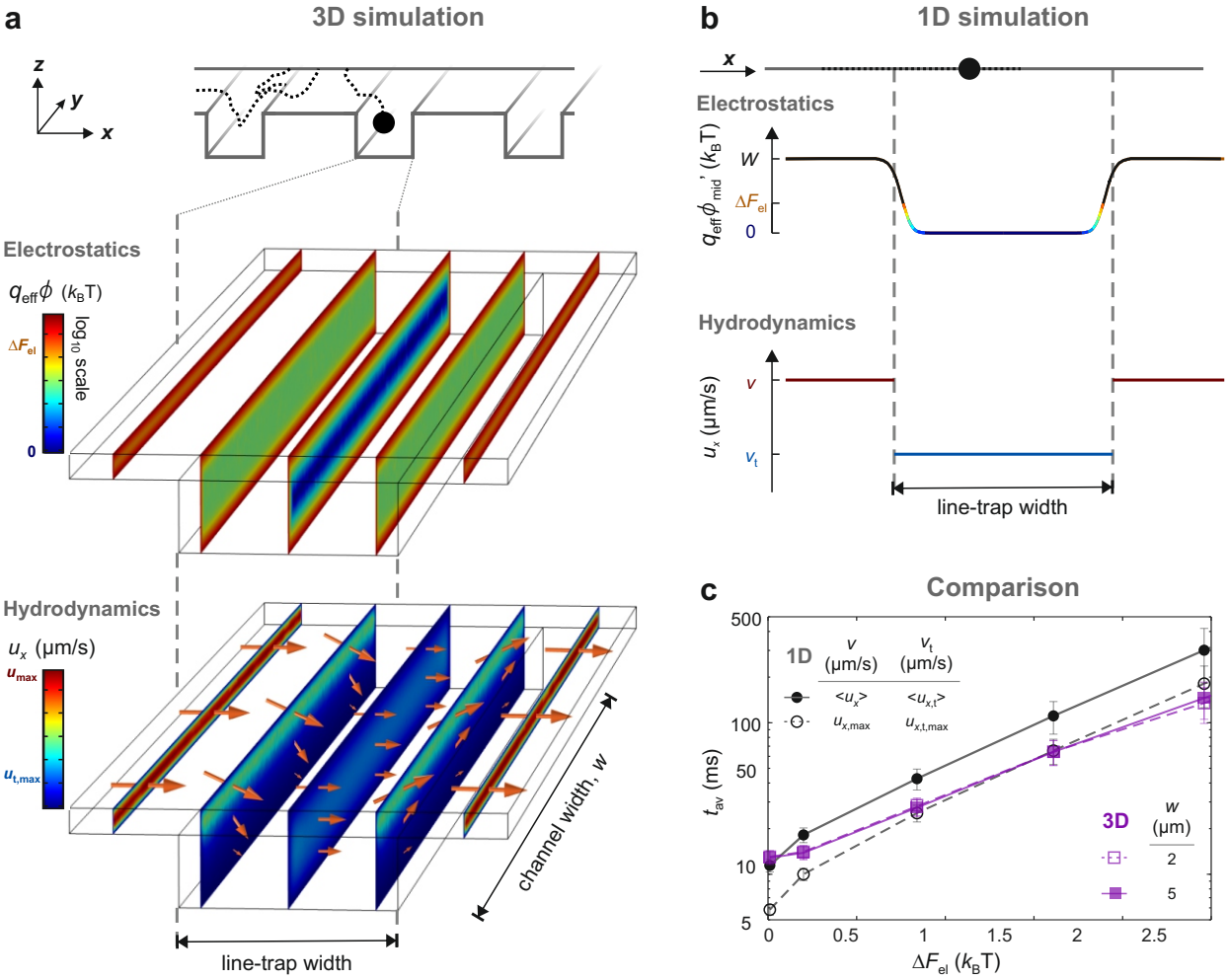


Fig. 25 **BD simulations of a FlET*e* experiment.** (**a**) 3D simulations were run of a particle diffusing in the FlET*e* trapping geometry superimposed with an electrostatic potential (middle panel) and velocity flow fields (bottom panel) as calculated using COMSOL. Electrostic potential, ϕ, was rescaled to give a variety of overall electrostatic well depths, ΔF_{el}, in simulations. Shown here is the COMSOL calculation of ϕ in 1 mM NaCl. Velocity flow fields, $\boldsymbol{u}$, were also rescaled to give a variety of centerline flow velocities along the channel, u_{max}, and corresponding velocities in the trap, $u_{\mathrm{t,max}}$. (**b**) 1D simulations were run of a particle diffusing along a line superimposed with an electrostatic potential and velocity flow field. The electrostatic potential shown is the sum of the ΔF_{el} from the 3D simulation and f, yielding the same total energy well depth W. The hydrodynamics imposed on the 1D simulation is a constant velocity flow, v or u_{t}, depending on if the molecule is in the channel or trap region respectively. (**c**) The escape times measured from the 1D simulations with $v{=}u_{\mathrm{max}}$ match well with the 3D simulations while $\Delta F_{\mathrm{el}} \gtrsim 1\ k_{\mathrm{B}}\mathrm{T}$, corresponding to $W \gtrsim 3k_{\mathrm{B}}\mathrm{T}$. As $\Delta F_{\mathrm{el}} \to 0$, the 3D simulations tend to the 1D simulation result with $v{=}\langle u\rangle$, as expected for decreasing midplane confinement of the particle.

The centre of the particle was defined by $\boldsymbol{r}(t)$ from the Langevin equation, and the finite size of the particle was explicitly accounted for by considering a shell of points at distance R from the centre, and ensuring that after each timestep none were outside the trapping geometry; the step was discarded upon such an occurrence. Δt and t_{esc} were extracted from the particle trajectories using escape thresholds of $R_{\mathrm{esc},1} = 200$ nm and $R_{\mathrm{esc},2} = 18718(R\mathrm{x}10^{9})^{-0.0037} - 18051 \approx 600$ nm for R expressed in m. Simulations were setup for a series of R, h_1 and h_2, spanning the range of physical parameters used in experiments in Chapter 4.

We envisage this simulation framework for 3D simulations of ETsy experiments can be used for further *in silico* studies of the escape dynamics of differently shaped particles, with explicit rotational diffusion and asymmetric diffusion all trivial to implement.

Appendix E

Experimental error in determining q_{eff}

When comparing measurements of a similar sample under different conditions, such as in section 2.2, we must consider only uncertainties that vary from measurement to measurement. This excludes uncertainty in D_{t}, ϕ_{s} and $2h$ for example, where inaccuracies are the same for a given sample in a given device from measurement to measurement. The main source of such uncertainty is from determining the escape time, whose error, $t_{\mathrm{esc,e}}$, is statistically limited by the number of escape events measured, N, as given by Eq. 25

$$t_{\mathrm{esc,e}} = \frac{t_{\mathrm{esc}}}{\sqrt{N}} \tag{25}$$

Additional uncertainties come from the salt concentration, c_{e}, and in the pH of a measurement, both of which determine the midplane potential of the device. Each of these are evaluated as the standard error of the mean from repeat measurements. In order to account for these we first express q_{eff} in terms of these parameters by combining Eqs. 10, 11 and 13, to give a final expression Eq. 26.

$$q_{\mathrm{eff}} = \left(k_{\mathrm{B}} T \ln\left(\frac{t_{\mathrm{esc}}}{t_{\mathrm{r}}}\right) - f \right) \frac{1}{2\phi_{\mathrm{s}}} \exp(\kappa h) \tag{26}$$

We note that for the measurements in this study κ approximately depends on c (expressed in M) by the following relation: $\kappa \approx \frac{\sqrt{c}}{0.304}$ nm. Thus the error on κ, κ_{e}, depends on c_{e} as follows: $\frac{\kappa_{\mathrm{e}}}{\kappa} = \frac{1}{2}\frac{c_{\mathrm{e}}}{c}$. We also note that due to the relatively weak dependence of ϕ_{s} on the pH (see Fig. 7), this error contributes negligibly to $q_{\mathrm{eff,e}}$.

We follow the general error propagation expression for the total error z_{e} on quantity $z(x,y)$ arising from errors in variables x and y given by Eq. 27:

$$\left(\frac{\mathrm{z_e}}{\mathrm{z}}\right)^2 = \frac{1}{\mathrm{z}^2}\left(\frac{\partial z}{\partial x}\right)^2 {x_{\mathrm{e}}}^2 + \frac{1}{\mathrm{z}^2}\left(\frac{\partial z}{\partial y}\right)^2 {y_{\mathrm{e}}}^2 \tag{27}$$

Using Eq. 27 we can express the uncertainty on q_{eff}, $q_{\mathrm{eff,e}}$, as Eq. 28.

$$\left(\frac{q_{\mathrm{eff,e}}}{q_{\mathrm{eff}}}\right)^2 = \frac{1}{{q_{\mathrm{eff}}}^2}\left(\frac{\partial q_{\mathrm{eff}}}{\partial t_{\mathrm{esc}}}\right)^2 {t_{\mathrm{esc,e}}}^2 + \frac{1}{{q_{\mathrm{eff}}}^2}\left(\frac{\partial q_{\mathrm{eff}}}{\partial \kappa}\right)^2 {\kappa_{\mathrm{e}}}^2 \tag{28}$$

By combining Eq. 25, 26 and 28 and accounting for the dependence of κ_{e} on c_{e}, we obtain a final expression as Eq. 29.

$$\frac{q_{\mathrm{eff,e}}}{q_{\mathrm{eff}}} = \frac{k_{\mathrm{B}}\,T}{\Delta F_{\mathrm{el}}}\frac{1}{\sqrt{N}} + \frac{\kappa\,h}{2}\frac{c_{\mathrm{e}}}{c} \tag{29}$$

We note that generally $c_{\mathrm{e}} < 0.03\,\mathrm{mM}$, and so the the first term dominates. With typical values of $\Delta F_{\mathrm{el}} \approx 2\,k_{\mathrm{B}}\,T$ and N=2000 events, fractional uncertainty in q_{eff} can be expected to be ca. 1%. Only in low salt conditions can $\frac{c_{\mathrm{e}}}{c} \lesssim 3\%$, where the second term in Eq. 29 can be significant. With values of $\kappa h \approx 2$, typical of measurements in section 2.2, the second term becomes ca. 3%.

Appendix F

The effect of SNR and photo-blinking on escape-time measurements

Escape times measured from low signal-to-noise ratio (SNR) experiments were observed to vary considerably when different imaging conditions were used. Specifically, the sampling frequency with which the movies were recorded had a large impact, with faster sampling yielding faster escape times. The fractional difference in t_{esc} with different sampling frequencies was larger with lower SNRs. Simulated movies reproduce the response of t_{esc} to sampling frequency when the experimental SNR is reconstructed (red curve vs blue curve in Fig. 26**a**). The simulations give us an insight into the cause of this effect, as we can dissect an identical particle trajectory reproduced under different sampling conditions and SNRs, as shown in Fig. 26**b**. We can see that in low SNR movies, occasionally we have a frame where the molecule signal is too low to be detected, as a result trapping events are sometimes split into two, effectively halving the Δt measured. This is more likely to occur for high sampling frequency movies when events are longer (as a number of frames), thus we anticipate that this effect is thus a function of sampling frequency to escape time ratio.

In order to compare the effects for experiments of vastly different t_{esc} (and also sampling frequencies), we rescale the sampling frequency to be measured in "frames per event", and normalise the escape times by t_{esc} measured at a sampling frequency corresponding to 2 frames per event. This then allows us to compile a series of simulations of different sized particles (D_{t}) in different well depths (W) and see if a consistent trend emerges. Such simulations, realised as movies of three different SNRs, are shown in Fig. 27. We find that the simulated data at a given SNR collapses onto a single curve when rescaled, giving credence to the theory of low SNR increasing

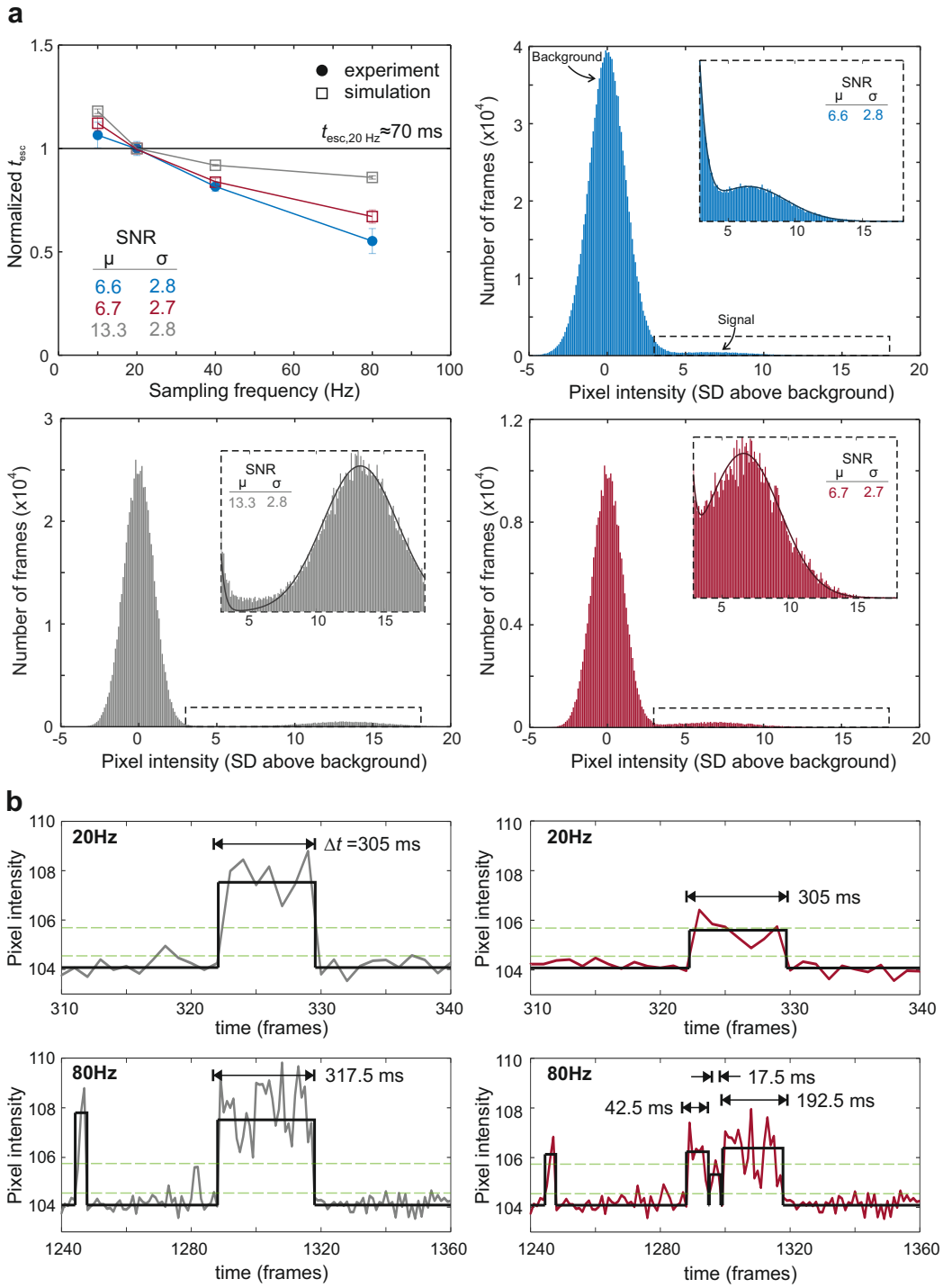


Fig. 26 **Effect of SNR on ET*e* experiments.** (**a**) A low SNR measurement of ATTO542-carboxyl at four different sampling frequencies (blue data); t_{esc} is plotted relative to the measurement at 20 Hz, and varies with changing sampling frequency. The SNR is measured as shown in right hand panel, with a two Gaussian fit to the distribution of pixel intensities (plotted as a function of standard deviations (SD) of the noise above the background) - one corresponding to the background, and one to molecule signal. Simulated movies of the experiment show the same trend when matching the SNR of the experiment (red data), while higher SNR simulated movies show a smaller variation with sampling frequency (grey data). (**b**) Intensity vs time plots from a pocket in simulated movies at high (left panels) and low SNRs (right panels), illustrating the cause of the sensitivity of t_{esc} to sampling frequency. Green horizontal lines illustrate the intensity thresholds applied to determine if a molecule is trapped or not.

the number of dropped frames during escape events, which affects the t_{esc} disproportionately for events many frames long. Two experimental datasets with vastly different escape times and SNR are overlayed and show qualitative agreeement with the simulations, although simulated movies generated at exactly the experimental SNRs would be required for complete agreement: ATTO542-carboxyl with $t_{esc} \approx 70$ms and SNR given by $\mu = 6.4$ and $\sigma = 2.7$ (orange data), and 20 nm fluorescent polystyrene beads with $t_{esc} \approx 850$ms and SNR given by $\mu = 13.4$ and $\sigma = 4.8$ (red data).

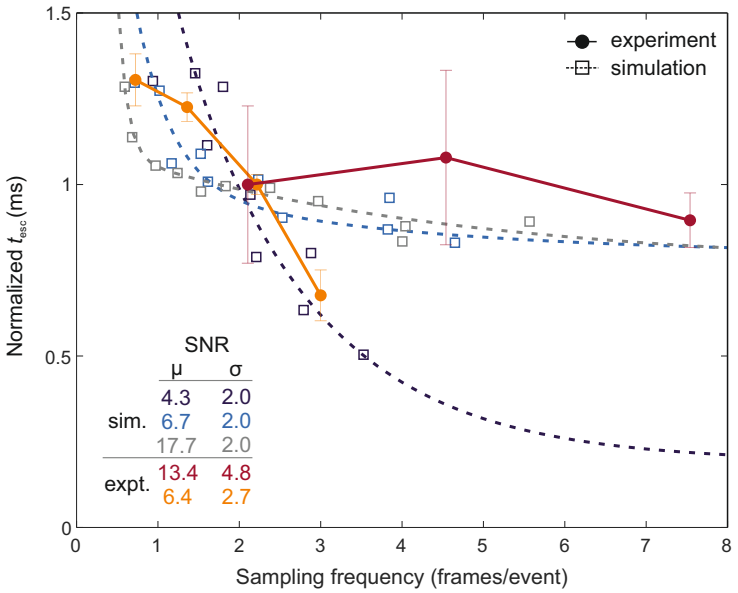


Fig. 27 **Global analysis of SNR influence on ET*e* experiments.** t_{esc} from simulated movies at different sampling frequencies for a given SNR fall on a single curve when correctly normalised (open circles and dashed lines). Experiments follow the same qualitative trend as these simulations: low SNR measurement of ATTO542-carboxyl (orange data) shows large variation in t_{esc} while a high SNR measurement of 20 nm fluorescent polystyrene beads (red data) shows no variation.

This plot highlights both the importance of simulated movies to correctly interpret t_{esc}, and also the importance of choice of sampling frequency in an experiment, both are especially important in low SNR experiments. However, we see that even high SNR movies show considerable variation in t_{esc} when the sampling frequency is $\lesssim 2$, so it is helpful to sample faster than this.

We note that the observation of undetected signal frames in the middle of trapping events due to low SNR in simulations, could also manifest in experiments due to photo-blinking; if a dye were to spontaneously stop emission for $\gtrsim$5 ms then there will be a dark frame likely to not be detected. This is similarly more likely to occur at high sampling frequency rates as a chance of a dark frame being in the middle of an event is increased. We again turn to simulations and impose a mask of dark frames on top of the simulated particle trajectory. The dark frames are imposed

to occur at a fixed rate, characterised by the timescale t_{on}, for a duration characterised by the timescale t_{off}. The dark frames are "invisible" to the analysis and thus if a particle is in a pocket during that frame the escape event will be split. Two different t_{on} rates of mask were applied to the same trajectory, and the effect is qualitatively the same as lowering the SNR of the movie - a larger variation in t_{esc} with sampling frequency is observed with more frequent dark frames (smaller t_{on}) see Fig. 28. Photo-blinking has been reported over timescales of milliseconds to hours, and thus the cases investigated here are entirely plausible[96].

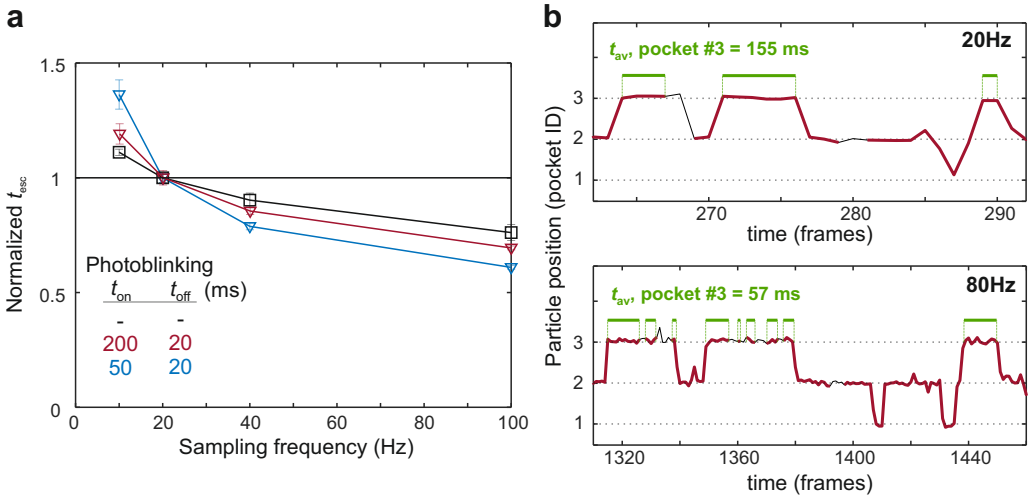


Fig. 28 **Effect of photo-blinking on ET*e* experiments.** (**a**) t_{esc} extracted from the same simulated particle trajectory but with different frequency dark-frame masks imposed, analysed at a variety of different sampling frequencies. With more frequent dark frames occurring at average interval t_{on} we see t_{esc} reduces more with sampling frequency (blue vs red curve) (**b**) Coordinate-time plots of the simulated trajectory (black line) with a dark frame mask imposed with $t_{\text{on}} = 200$ ms, $t_{\text{off}} = 20$ ms (red line), sampled at 20 Hz (top panel) and 80 Hz (bottom channel). Measured Δt in pocket #3 are explicitly shown (green lines).

Appendix G

Supporting information for "Far-field electrostatic signatures of macromolecular 3D conformation"

Supporting Information for

Far-field electrostatic signatures of macromolecular 3D conformation

Gunnar Kloes[†1], *Timothy J. D. Bennett*[†1], *Alma Chapet-Batlle*[2], *Ali Behjatian*[1], *Andrew J. Turberfield*[2] *& Madhavi Krishnan**[1]

[1]Physical and Theoretical Chemistry Laboratory, Department of Chemistry, University of Oxford, South Parks Road, Oxford OX1 3QZ, United Kingdom

[2]Clarendon Laboratory, Department of Physics, University of Oxford, Parks Road, Oxford OX1 3PU, United Kingdom

*Corresponding author. Email: madhavi.krishnan@chem.ox.ac.uk

[†]these authors contributed equally to the work

Table of contents:

1. **DNA Nanostructure synthesis and characterisation……2**

   Figures S1-5

   Tables S1-5

2. **Escape-time electrometry (ET*e*) to measure the effective charge, q_m, of DNA nanostructures……15**

   Figures S6,7

3. **Procedure to calculate values of effective charge, q_{eff}, for DNA nanostructures……30**

   Figures S8,9

Supporting References……...37

1. DNA Nanostructure synthesis and characterization

1.1: Nanostructure characterization and general procedure for synthesis

Sequences. Oligonucleotide sequences are recorded in Tables S1-4.

Oligonucleotide synthesis. DNA oligonucleotides were ordered from Integrated DNA Technologies, Inc (IDT), unpurified except for those with the 5'Atto-532 fluorescent label (HPLC purified). Strands were resuspended at a concentration of 100μM in TE Buffer (10mM Tris + 1mM EDTA).

Assembly. Annealing protocols for each nanostructure are specified separately below.

Gel characterization. Nanostructures were characterized by polyacrylamide gel electrophoresis (PAGE) using 7.5% polyacrylamide gels (29:1 acrylamide : bisacrylamide) run in 1× TAE buffer (40mM Tris, 20mM acetic acid, 1mM EDTA) at 180 V for 35 min. Gels presented in Figures S1-5 were imaged in the Atto532 fluorescence channel (532nm laser and 570nm bandpass filter) with an Amersham Typhoon Gel Imaging System (GE Healthcare).

Gel purification. Gels (run as above) were imaged in the Atto532 fluorescence channel (532nm laser and 570nm bandpass filter). The band corresponding to a fully-assembled nanostructure was cut out and nanostructures eluted by adding 200 μL TE Buffer with 5 mM $MgCl_2$ (tetrahedron, two-helix bundle, 120bp DNA helix) or 12.5 mM $MgCl_2$ (square tile), freezing and thawing twice and collecting the supernatant. Tetrahedron and square tile samples were further concentrated by ethanol precipitation[1] and resuspended in TE Buffer alone (square tile) or in TE Buffer with 5 mM MgCl2 (tetrahedron).

Quantification. Purified sample concentrations were estimated on the basis of the intensity of emission from the Atto532 fluorophore. Fluorescence intensities were recorded using a CLARIOstar Plus (BMG Labtech) Microplate reader (excitation: 520nm, emission: 570nm). A calibration curve was obtained using the 5'Atto-532 labeled strand T_s1 at known

concentrations ranging from 100 pM to 100 nM. Measurements were performed in triplicate and averaged.

1.2: Tetrahedron synthesis

Design. The tetrahedron was synthesized as described in Goodman *et al.* [2] with the addition of an Atto532 fluorescent label on strand T_s2 (Figure S1a).

Tetrahedron Assembly. Labeled strand (T_s2) at 50nM and unlabeled strands (T_s1, T_s3, T_s4) at 75 nM were combined in TE Buffer with 5 mM MgCl. The mixture was heated to 95°C, held for one minute, then cooled at 2°C per minute to 22°C.

1.3: Square tile synthesis

Design. The tile comprises four DNA double-helices, each of length of 30 bp, bound to neighboring helices through crossovers at the helix ends. Nicks (Figure S2a,b, helix 1, 2 and 3) were placed asymmetrically, dividing a helix into 20bp and 10bp domains, in order to reduce the potential for aggregation during assembly.

Assembly. Labeled strand (Sq_s1) at 50 nM and unlabeled strands (Sq_s2, Sq_s3, Sq_s4, Sq_s5) at 75 nM were combined in TE Buffer with 12.5 mM MgCl. The mixture was heated to 95°C, held for one minute, then cooled at 1°C per minute to 22°C.

1.4: Bundle synthesis

Design. The two-helix bundle is a double crossover tile[3] comprising two 60-nucleotide DNA helices bound together by two crossovers. It is assembled from five strands of DNA, two that run the full length of the helices and three that bind these two strands together (Figure S3a,b). There is a single nick on one of the helices where the 5' and 3' ends of the central linking strand meet. The spacing between the central crossovers was chosen to be 21 bp, corresponding to two turns of the DNA double-helix.

Assembly. Labeled strand (S2) at 200 nM and unlabeled strands (Figure S3, B_s1, B_s2, B_s3) at 300nM were combined in TE Buffer with 12.5 mM MgCl. The mixture was heated to 95°C, held for one minute, then cooled at 1°C per minute to 22°C.

1.5: 120bp DNA helix synthesis

Design. The 120bp helix is assembled from four strands of DNA to circumvent the requirement to synthesize an Atto532-labeled oligonucleotide of length 120 nucleotides (Figure S4a). Strands S1 and S2 from the 2-helix bundle are re-used in this structure.

Assembly. Labeled strand (S1) at 200nM and unlabeled strands (S2, Ds_s1, Ds_s2) at 300 nM were combined in TE Buffer with 12.5 mM MgCl. The mixture was heated to 95°C, held for one minute, then cooled at 1°C per minute to 22°C.

1.6: Determining the structural charge of DNA nanostructures

The structural charge, q_{str}, is determined for each of the DNA nanostructures as follows. Beyond summing the number of bases, n_{b} (each associated with a charge of -1e due to a phosphate ester group), we also consider the number of strands, N_{s}, in each structure, and whether their 5' termini are phosphorylated or not. In the case of all the nanostructures, ($N_{\mathrm{s}} - 1$) strands had hydroxylated 5' termini, while the dye-carrying "N_{s}th strand" did carry a 5' phosphorylation to which the dye is chemically coupled. The 60bp double-stranded calibrator molecule had both strands phosphorylated at the 5' end. We also add a contribution of -1e to the total structural charge for each Atto-532 dye attached to the DNA nanostructure, as quoted on the manufacturer's website (https://www.atto-tec.com/support/datatable/) and independently verified in previous ETe measurements[4]. Table S5 summarizes the charge calculation.

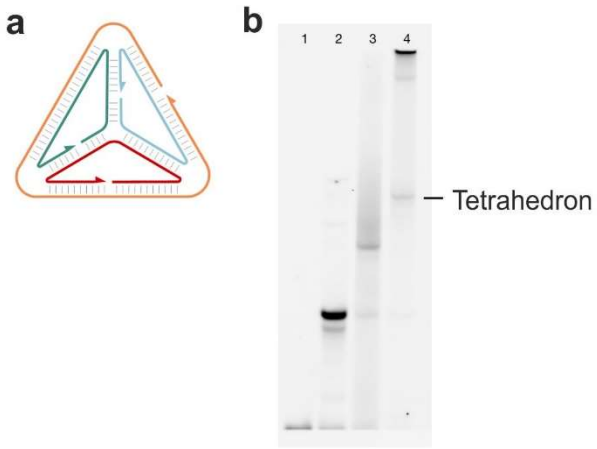


Figure S1. Design & assembly of the tetrahedron. (**a**) Tetrahedron (Figure adapted from Goodman *et al.* [2]) (**b**) Formation gel: each lane contains the unpurified products of assembly of one more strand than the previous lane. Lane 1: T_s1; Lane 2: T_s1 + T_s2; Lane 3: T_s1 + T_s2 + T_s3; Lane 4: T_s1 + T_s2 + T_s3 + T_s4.

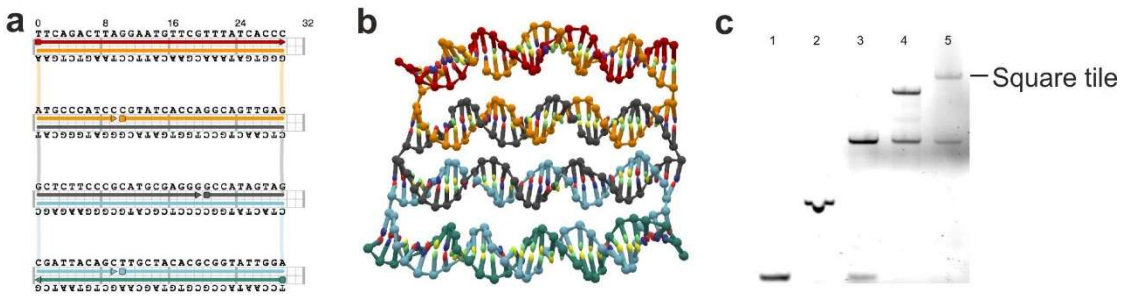


Figure S2. Design & assembly of the square tile. (**a**) scadnano[5] design of the square tile. (**b**) OxDNA[6-8] representation of the square tile (**c**) Formation gel: each lane contains the unpurified products of assembly of one more strand than the previous lane. Lane 1: Sq_s1; Lane 2: Sq_s1 + Sq_s4; Lane 3: Sq_s1 + Sq_s4 + Sq_s2; Lane 4: Sq_s1 + Sq_s4 + Sq_s2 + Sq_s5; Lane 5: Sq_s1 + Sq_s4 + Sq_s2 + Sq_s5 + Sq_s3.

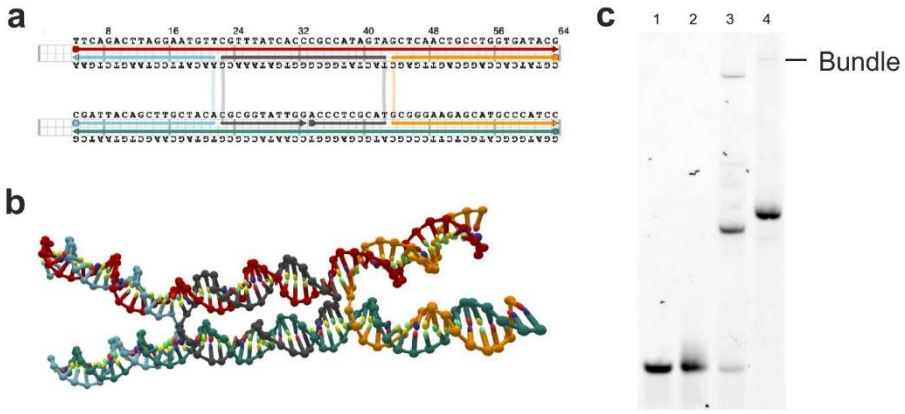


Figure S3. Design & assembly of the two-helix bundle. (**a**) scadnano[5] design of the Two-Helix Bundle. (**b**) OxDNA[6-8] representation of the designed Two-Helix Bundle (**c**) Formation gel: each lane contains the unpurified products of assembly of one more strand than the previous lane. Lane 1: S1; Lane 2: S1 + B_s2; Lane 3: S1 + B_s2 + S2; Lane 4: S1 + B_s2 + S2 + B_s1; Lane 5: S1 + B_s2 + S2 + B_s1 + B_s3

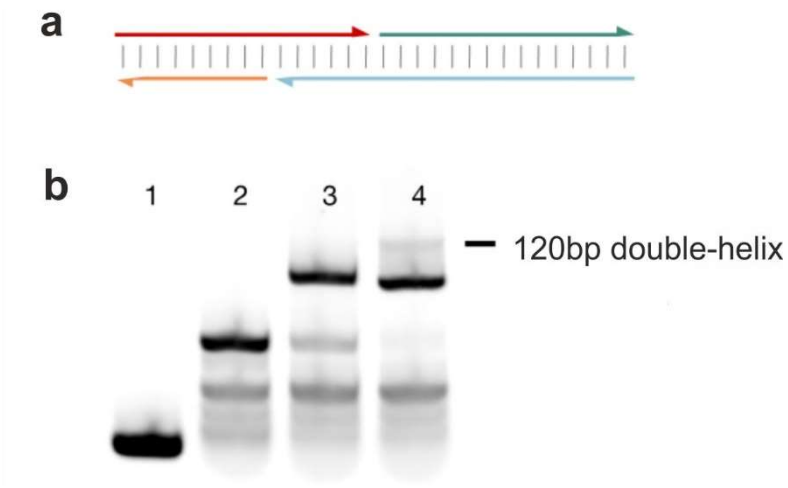


Figure S4. Design & assembly of the 120bp DNA helix. (**a**) 120 bp DNA Helix (**b**) Formation gel: each lane contains the unpurified products of assembly of one more strand than the previous lane. Lane 1: S1; Lane 2: S1 + Ds_s2; Lane 3: S1 + Ds_s2 + S2; Lane 4: S1 + Ds_s2 + S2+ Ds_s1.

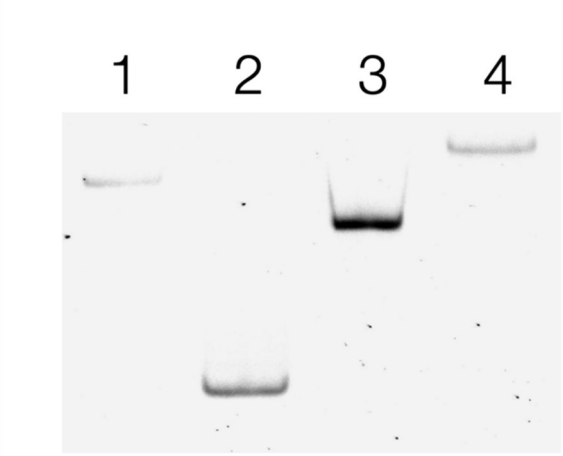


Figure S5. Native PAGE of all four gel purified structures imaged in the Atto532 fluorescence channel. Lane 1: Double-stranded DNA; Lane 2: Two-Helix Bundle; Lane 3: Square Tile; Lane 4: Tetrahedron.

Table S1. Tetrahedron strand sequences. All sequences are written 5' to 3'.

T_s1	AGGCAGTTGAGACGAACATTCCTAAGTCTGAAATTTATCACCCGCCATAGTA GACGTATCACC
T_s2	(5'ATTO532)CTTGCTACACGATTCAGACTTAGGAATGTTCGACATGCGAGGG TCCAATACCGACGATTACAG
T_s3	GGTGATAAAACGTGTAGCAAGCTGTAATCGACGGGAAGAGCATGCCCATCC ACTACTATGGCG
T_s4	CCTCGCATGACTCAACTGCCTGGTGATACGAGGATGGGCATGCTCTTCCCGA CGGTATTGGAC

Table S2. Square Tile strand sequences. All sequences are written 5' to 3'.

Sq_s1	(5'ATTO 532) TTCAGACTTAGGAATGTTCGTTTATCACCC
Sq_s2	GCCATAGTAGCTCAACTGCCTGGTGATACGGGATGGGCATGCTCTTCCCGCA TGCGAGGG
Sq_s3	TCCAATACCGCGTGTAGCAAGCTGTAATCG
Sq_s4	CGTATCACCAGGCAGTTGAGGGGTGATAAACGAACATTCCTAAGTCTGAAA TGCCCATCC
Sq_s5	TTGCTACACGCGGTATTGGACTACTATGGCCCCTCGCATGCGGGAAGAGCCG ATTACAGC

Table S3. Two-Helix Bundle strand sequences. All sequences are written 5' to 3'.

S1	(5'ATTO532)TTCAGACTTAGGAATGTTCGTTTATCACCCGCCATAGTAGCTCA ACTGCCTGGTGATACG
S2	GGATGGGCATGCTCTTCCCGCATGCGAGGGTCCAATACCGCGTGTAGCAAG CTGTAATCG
B_s1	CGTATCACCAGGCAGTTGAGCGCGGGAAGAGCATGCCCATCC
B_s2	ACCCTCGCATTACTATGGCGGGTGATAAACGCGCGGTATTGG
B_s3	CGATTACAGCTTGCTACAAACATTCCTAAGTCTGAA

Table S4. 120bp DNA helix strand sequences. All sequences are written 5' to 3'.

S1	(5'ATTO532)TTCAGACTTAGGAATGTTCGTTTATCACCCGCCATAGTAGCTCA ACTGCCTGGTGATACG
S2	GGATGGGCATGCTCTTCCCGCATGCGAGGGTCCAATACCGCGTGTAGCAAG CTGTAATCG
Ds_s1	GCATGCGGGAAGAGCATGCCCATCCCGTATCACCAGGCAGTTGAGCTACTA TGGCGGGTGATAAACGAACATTCCTAAGTCGTAA
Ds_s2	CGATTACAGCTTGCTACACGCGGTATTGGACCCTC

Table S5. Physical properties of DNA nanostructures. Nanostructure properties listed in the table are as follows: the longest nanostructure dimension, l, calculated assuming 0.34 nm per base pair; total number of bases, n_{b}; number of individual single strands in the nanostructure, N_{s}; number of Atto532 fluorophores, N_{dye}; structural charge, q_{str}, calculated for each nanostructure using $-[n_{\mathrm{b}} + (N_{\mathrm{s}}\text{-}1) + N_{\mathrm{dye}}]$, and for the calibrator molecule using $-[n_{\mathrm{b}} + N_{\mathrm{dye}}]$.

Nanostructure	l (nm)	n_{b}	N_{s}	N_{dye}	q_{str} (e)
60ds rod (calibrator)	20.4	120	2	2	-122
120ds rod	40.8	240	4	1	-238
bundle	20.4	240	5	1	-237
tile	10.4	240	5	1	-237
tetrahedron	5.44	252	4	1	-250

2. Escape-time electrometry (ET*e*) to measure the effective charge, q_{m}, of DNA nanostructures

Overview of method

We use ET*e* to measure electrostatic interaction energies and therefore the effective electrical charge of molecules in solution, with high precision. Briefly, the method relies on optical observation of charged molecules diffusing in an electrostatic interaction energy landscape generated in a fluid-filled gap between parallel like-charged surfaces (Figure S6a). We measure the reduction in system free energy associated with transferring an electrically charged object from regions in which the parallel-plates are closely spaced into nanostructured 'trap' regions of very weak confinement. At the location of a trap, the two flanking surfaces are much further apart than in the slit region, and the molecule-plate electrical repulsion here is therefore negligible (Figure S6c). We create an array of such electrostatic fluidic traps using periodic nanoscale indentations in the upper plate of a slit created by two parallel silica surfaces, where the typical height of the gap is $2h = 70$ nm (Figure S6a).

Using wide-field fluorescence microscopy we observe the diffusive migration of molecules in a landscape of effectively identical electrostatic traps. We measure an average residence time, or escape time from the trap, t_{esc}, of the molecular species of interest. The escape time yields a sensitive measure of the depth of the potential well or the trap depth, W, created by the nanostructured surface indentations, which in turn depends strongly on the effective charge of the molecular species. In order to determine the effective charge we first infer a value of potential well depth, W, from our measurement of t_{esc}, and then determine the electrostatic contribution, ΔF_{el}, to the total trap depth, in order to infer a value of the measured effective charge, q_{m}.The experimentally inferred effective charges are denoted as q_{m} to distinguish a measured quantity from a theoretically calculated quantity, q_{eff} . Note that due to

spatial and rotational fluctuations of the molecule, we have $W > \Delta F_{\mathrm{el}}$ in general, described in detail in section 2.4.

2.1: Measuring the average species escape times, t_{esc}, from an electrostatic fluidic trap.

DNA nanostructures were stored at a concentration of approximately 5-15 nM in TE buffer (10 mM Tris (tris(hydroxymethyl)aminomethane), 1 mM EDTA (Ethylenediaminetetraacetic acid) and 5 mM $MgCl_2$ (except the tile conformation which was stored with no $MgCl_2$). Immediately prior to measurement DNA nanostructures were diluted to a final concentration of 5-10 pM in buffer solutions of 1 mM Tris, ≈0.005x TE and ≈5 μM $MgCl_2$, with the final ionic strength, I, tuned by adding NaCl. We parametrize experimental conditions using the dimensionless product, κh, of inverse electrostatic screening length, κ (inverse Debye length), and h, which denotes half the height of the slit. κ is obtained from electrical conductivity measurements of the measured aqueous suspension in the device, after each ETe measurement (HORIBA LAQUAtwin), as described previously[9]. The slit height, $2h$, is characterized using atomic force microscopy (Asylum Research); the device fabrication procedure is similar to that in previous work[10].

For a single measurement of a nanostructure species at a specific ionic strength, molecules are loaded into nanoslits housing several arrays of traps using a vacuum suction system. Once loaded, the vacuum is turned off initiating free diffusion of the DNA nanostructures across the trap landscape. Fluorescent signals from individual nanostructures were recorded using a standard wide-field fluorescence microscope, imaged using 5ms exposure times at frame rates from 5 to 50 Hz for up to 1000 frames on a sCMOS camera (Prime 95B, Photometrics). More molecules were then loaded into the trap arrays, and the procedure repeated about 20 times per measurement. From these videos of DNA nanostructure diffusion in the electrostatic free energy landscape, molecular residence times, Δt, in each trap were extracted using customized video analysis software. The average escape time, t_{esc}, for

each measurement was determined by fitting the probability density distribution, $P(\Delta t)$, with a function the form $P(\Delta t) = \frac{A}{t_{\mathrm{esc}}} \exp(-\frac{\Delta t}{t_{\mathrm{esc}}})$, as described in previous work [4,11,12]. Previous studies confirm that photobleaching does not affect the measurement of escape times in our experiment[9].

2.2: Brownian dynamics (BD) simulations of escape process to convert the measured escape times, t_{esc}, to well depth, W, in the free energy landscape.

We converted measurements of escape time to potential well depth values, W, as described previously[4,9,11,12]. As a general principle, when $W > 4\,k_{\mathrm{B}}T$, we expect an exponential dependence of t_{esc} on W of the form

$$t_{\mathrm{esc}} = t_{\mathrm{r}} \exp(W/k_{\mathrm{B}}T). \quad \text{(S1)}$$

Here, t_{r} is an effective molecular position relaxation time which is inversely related to the molecular diffusion coefficient, D, i.e., $t_{\mathrm{r}} \propto \frac{1}{D}$. While equation S1 holds for large well depths and provides qualitative insight into the underlying dynamics, it does not strictly hold for well depths $W < 4\,k_{\mathrm{B}}T$. In practice we do not know *a priori* the value of W for a given molecular species. We therefore perform Brownian Dynamics (BD) simulations of the escape process in order to accurately relate t_{esc} to W as described previously[9,12]. We simulate the thermal motion of a charged object in an electrostatic energy landscape, where the object is characterised by a friction coefficient given by $6\pi v \eta r_{\mathrm{H}}$, while moving at velocity, v, through a medium of dynamic viscosity, η. The hydrodynamic radius, r_{H}, of the molecular species is determined using two-focus Fluorescence Correlation Spectroscopy (2fFCS) to measure the molecular diffusion coefficient, which in turn may be expressed in terms of the radius of an equivalent sphere, r_{H} (see section 2.3). The electrostatic potential landscape in the fluidic nanostructure is calculated by numerically solving the Poisson-Boltzmann (PB) equation, for our device geometry, subject to the appropriate boundary conditions, using COMSOL

Multiphysics as described previously[9,12]. The electrostatic interaction energy of the particle at any point $\mathbf{r}$ in the landscape, referenced to a free energy at infinite slit height, is given by $F_{\mathrm{el}} = q_{\mathrm{eff}}\phi(\mathbf{r})$, where $\phi(\mathbf{r})$ is the electrical potential at $\mathbf{r}$. A periodic array of axially symmetric potential wells, constructed from the minimum free energy profile of a charged object traversing a single trap, is used for subsequent BD simulations (main text Figure 1c).

Simulated trajectories of particle motion in the landscape are used to generate image stacks similar to the experimental movies, as described previously[12]. Briefly, simulated movies are analysed in the same way as experimental movies in order to extract escape times (or average residence times of particles at trap locations) as a function of known input well depths, W. While the generation of simulated movies fosters high precision measurements (<1% on q_{m}), movie generation is computationally expensive and not necessary for measurements with lower accuracy requirements (uncertainty around 5%). Here, we work directly with the spatial coordinates corresponding to the trajectory of particle motion in the landscape as described in detail previously[9]. We apply spatial entry and exit criteria to the simulated particle trajectory in order to define a trapping event. We point out that the molecular entry and exit criteria required in the trajectory-based analysis are initially determined by a tuning procedure which aims to recover the same escape time in both the simulated trajectory-based analysis and in the simulated movies. Here we first generate simulated movies of the escape process using the simulated molecular trajectories and the same signal-to-noise and brightness characteristics as the experiment. We then analyse the simulated movies in the same way as the experimental recordings. The entry and exit criteria in the trajectory-based analysis are fine-tuned until we obtain excellent agreement between average escape times from the particle trajectory and the simulated movie analyses. Thereafter we use just the simulated trajectory data, in conjunction with the entry and exit criteria, in order to theoretically determine escape time values that correspond to known input well depths, W. By comparing experimentally measured escape

times with theoretically expected escape times determined for a range of well depths, W, for a particle of known r_{H}, we determine the value of the unknown well depth in the experiment.

We see from equation S1 that the escape time depends exponentially on W and hence on the effective charge of the molecule, but only linearly on its size. Therefore our measured q_{m} values do not depend sensitively on accurate knowledge of molecular size, and a reasonable estimate of r_{H} will generally suffice. Note that as W always remains unknown at the outset, we do not *a priori* use equation S1 to convert t_{esc} to W. Once a set of simulations is performed for various value of W, we generally find that an exponential function of the form shown in equation S1 may be fit in the high well depth regime. For the DNA nanostructures in the study, we find $t_{\mathrm{r}} \approx 1.9$ ms, while for 60bp DNA, $t_{\mathrm{r}} \approx 1.4$ ms, under the experimental sampling conditions.

2.3: Measurement of hydrodynamic radii, r_{H}, of DNA nanostructures.

Dual-focus fluorescence correlation spectroscopy (2fFCS) was used to determine the effective hydrodynamic radii, r_{H}, of the DNA nanostructures. A homebuilt dual-focus confocal epi-fluorescence microscope was used to record fluorescence signals from labeled nanostructures freely diffusing in solution. Temporal correlation of the signals from the two foci yields a characteristic timescale for the diffusive transport of a molecular species between the two optical foci, from which the hydrodynamic radius can be inferred.

FCS measurements were performed on DNA nanostructures at 1-10 nM concentration in a buffer containing 10 mM Tris, 1 mM EDTA and 5 mM $MgCl_2$ (no $MgCl_2$ was used for measuring the tile sample). The molecules are excited by 5ps pulses of 520±15 nm laser light (NKT Photonics EXB-6 supercontinuum white light source, in conjunction with an NKT Photonics superK Varia wavelength selector) at 25 ns intervals alternatively at two confocal foci, spatially separated by a known distance. The pulses at 20 MHz are separated by a polarizing beam-splitter (Thorlabs), with a 25 ns temporal delay imposed on one arm via ~7.5

m of optical fiber. A Nomarski prism (Olympus) before the objective (1.2NA 60x Olympus) is then used to recombine the beams such that the two polarizations are brought to separate foci separated by a well-defined distance. The polarization selected time tagged time-resolved (TTTR) fluorescent signal is recorded on two avalanche photodiode single photon detectors (Micro Photon Devices, band-passed using filters from Chroma, Thorlabs and Semrock), and the cross-correlation function (CCF) between the TTTR signals is calculated (Figure S6a). Cross-correlation yields a timescale for diffusion between the optical foci, and hence a characteristic diffusion coefficient, D, for the species as outlined by Dertinger *et al.*[13]. The hydrodynamic radius, r_{H}, is calculated from D using the Stokes-Einstein equation $r_{\mathrm{H}} = k_{\mathrm{B}}T/6\pi\eta D$, with η the viscosity of water at temperature, T (Figure S6b).

We note that despite a significant disparity in the largest characteristic dimension of the nanostructures (compare a length of 41 nm for the cylindrical 120 bp DNA with ca. 8 nm for the side of the tetrahedron), and a corresponding variation of about a factor of 3 in the radius of gyration, R_{g}, the hydrodynamic radius has approximately the same value of $r_{\mathrm{H}} \approx 6 - 7$ nm for all DNA nanostructures (Figure S6b). We also point out that for all the DNA nanostructures in this study the electrostatic R_{g} turns out to be nearly identical to the standard definition of radius of gyration based on atomic positions in the molecule.

Measurement uncertainty in r_{H} is attributed to both measurement uncertainty in the diffusion coefficient, D_{e}, and in temperature, T_{e}, which influences r_{H} both explicitly through the Stokes-Einstein equation and also implicitly through its effect on viscosity, η. D_{e} is quantified by statistical bootstrapping which involves averaging over several fitting repetitions

of the CCF, each with a different random subset of all recorded data. T_e, is determined by the precision of the lab thermometer. Both uncertainties are accounted for in the quoted error, $r_{H,e}$.

2.4: Extraction of measured values of effective charge, q_m, from W

Having converted the measured escape times to potential well depth values, W, we next determine the electrostatic interaction free energies, ΔF_{el}, from which we deduce measured effective charge values, q_m, as described in previous work and summarized briefly here[4,9]. The well depth, W, can be broken down into the sum of two contributions: the free energy difference, ΔF_{el}, for states corresponding to electrostatic energy minima when the object is located at the midplane of the slit and the pocket regions, and a fluctuation contribution, f, that accounts for the difference in free energy due to axial and rotational motion of the object at each location[4,9,12,14]. ΔF_{el} is the difference in electrostatic free energy of interaction, $\Delta F_{el} = F_{el}|_{slit} - F_{el}|_{pocket} = q_{eff}(\phi_{mid}|_{slit} - \phi_{mid}|_{pocket})$, for a molecule positioned at the local energy minimum in the slit region, referenced to that in the pocket region[9,14]. In the absence of spatial fluctuations we would have $f = 0$ and therefore $W = \Delta F_{el}$ But since thermal fluctuations are always present, we write in general

$$W = W|_1 - W|_2 = \Delta F_{el} + f. \quad \text{(S2)}$$

F_{el} has robust theoretical underpinnings in the Poisson-Boltzmann (PB) framework for solution phase electrostatics, and can be calculated by first solving the non-linear PB equation,

$$\nabla^2\psi = \kappa^2 \sinh\psi \quad \text{(S3)}$$

subject to the appropriate boundary conditions for the experimental system as described previously[14,15] (see section 3 for further detail). Here $\psi = e\phi/k_B T$ is the dimensionless local electrical potential, ϕ. We have previously shown that ΔF_{el} is given by the product of an effective charge of the molecule in solution, q_{eff}, and the difference in midplane electrical potential, $\Delta\phi_{mid}$, between the slit and nanoscale pocket regions[14] (main text Figure 1).

Denoting the electrical potential at the midpoint between the parallel-plate slit surfaces as ϕ_{m}, and recognizing that $\phi_{\mathrm{mid}} = 0$ by design in the pocket region, we therefore have

$$\Delta F_{\mathrm{el}} = q_{\mathrm{eff}}\phi_{\mathrm{m}} \quad \text{(S4)}$$

We emphasize that the relationship $F_{\mathrm{el}} = q_{\mathrm{eff}}\phi(\mathrm{r})$, on which Eq. S4 is based, has been shown to hold rigorously in theoretical treatments such as Kjellander's "dressed ion" theory[16,17].

Thus in order to obtain a measure of ΔF_{el} and hence of q_{eff} from the inferred well depth, *W*, we first need to determine the fluctuation contribution, *f*, to the total well depth, W. The dominant contribution to the fluctuation free energy comes from axial fluctuation of the molecule in the z-dimension in the slit as described previously[4,14]. There is also a small contribution from a rotational term which we consider for non-spherically symmetric objects. We determine $f = \Delta F_{\mathrm{trans}} + \Delta F_{\mathrm{rot}}$ by considering the differences in an 'excess' free energy due to translational ($\Delta F_{\mathrm{trans}}$) and rotational ($\Delta F_{\mathrm{rot}}$) fluctuations of the molecule between the two states of our parallel plate system. $\Delta F_{\mathrm{trans}}$ accounts for the thermal motion of a molecule about the minimum-energy midplane of the system. As shown previously, we can write $\Delta F_{\mathrm{trans}}$ in terms of the partition function, *Z*, of a point charge diffusing freely in *z* as follows[4]:

$$\Delta F_{\mathrm{trans}} = -k_{\mathrm{B}}T\ln Z|_1 + k_{\mathrm{B}}T\ln Z|_2 - q_{\mathrm{eff}}\phi_{\mathrm{m}} \quad \text{(S5)}$$

$$Z = \int_0^{z_{\max}} e^{-\Delta F_{\mathrm{el}}(z)/k_B T}\, dz \quad \text{(S6)}$$

where $\Delta F_{\mathrm{el}}(z) = q_{\mathrm{eff}}\phi(z)$ (S7)

represents the electrostatic free energy of the charged object located at any axial position, *z*, in the slit (referenced to the global minimum in interaction energy ($z \to \infty$)).

Solving the equation S3 for a monovalent, binary and symmetric electrolyte, subject to constant charge boundary conditions, shows that the electrical potential at a location *z* in a parallel plate slit can be approximated as $\phi(z) = \phi_{\mathrm{mid}}\cosh\left[\kappa\left(z - \frac{z_{\max}}{2}\right)\right]$.

Using equations S6 and S7, we evaluate *Z* for particle states 1 and 2 (the "slit" and "pocket" states, see Fig. S8a), with upper limits on the integral given by $z_{\max} = 2h$ or

$(2h + d)$, dictated by the maximum parallel plate separation in the slit $(2h)$ or in the pocket region $(2h + d)$. Note that $\phi(z_{\max}/2) = \phi_{\mathrm{mid}}$ is the electrical potential at the midplane of the slit.

ΔF_{rot}, in turn, accounts for the additional free energy due to rotational motion of a non-spherical particle. By approximating the DNA nanostructure as a line charge of length, l, given by the longest dimension of the nanostructure, carrying a charge density $\sigma_l = q_{\mathrm{eff}}/l$, we can write ΔF_{rot} in terms of the partition function, Ω, of a freely rotating rod at an angle, θ, to the z-axis, as follows:

$$\Delta F_{\mathrm{rot}} = -k_{\mathrm{B}}T\ln\Omega|_1 + k_{\mathrm{B}}T\ln\Omega|_2 - q_{\mathrm{eff}}\phi_{\mathrm{m}} \quad \text{(S8)}$$

Here we write the following expression for the local angular electrostatic energy, $\Delta F_{\mathrm{rot}}(z,\theta)$, of a rod of length l located at a height z far away from the slit walls, oriented at an angle θ to the z-axis

$$\Delta F_{\mathrm{rot}}(z,\theta) = \frac{1}{\cos\theta}\int_{z-\frac{l}{2}\cos\theta}^{z+\frac{l}{2}\cos\theta} \sigma_l\phi(z)dz \quad \text{(S9)}$$

We may also write the above equation in simpler form: $\Delta F_{\mathrm{rot}}(z,\theta) = q_{\mathrm{eff}}\langle\phi(z)\rangle$, where $\langle\phi(z)\rangle$ denotes the average electrical potential evaluated over a height in the slit ranging from $z - \frac{l}{2}\cos\theta$ to $z + \frac{l}{2}\cos\theta$. For example, for $\theta = 90°$, and $z = z_{\max}/2$, equation S9 reduces to $\Delta F_{\mathrm{rot}}(z,\theta) = q_{\mathrm{eff}}\phi_{\mathrm{mid}}$, which is the electrostatic interaction energy of a rod of charge q_{eff} lying along the mid-plane of the slit. Since the problem is azimuthally symmetric around the z-axis of the slit, the angular partition function for a rod centred on and confined to the midplane of the parallel-plate gap at $z = z_{\max}/2$ is given by

$$\Omega = 4\pi\int_0^{\pi/2} e^{-\Delta F_{\mathrm{rot}}(z_{\max}/2,\theta)/k_{\mathrm{B}}T}\sin\theta\, d\theta \quad \text{(S10)}$$

In order to simplify the calculation of the rotational term, we have assumed that the charged object is confined to the midplane of the system, i.e. we only consider rotational contributions for the molecule located at $z = z_{\max}/2$. The accuracy of this approximation, which permits

greatly simplified calculation of ΔF_{rot}, was verified by comparison with an explicit full translational-rotational free energy calculation.

When calculating ΔF_{rot} for the different DNA nanostructures, we consider rods of different length, reflecting the longest dimension of the nanostructure. We assume a length of 0.34 nm per base pair which gives nanostructure lengths, l, as shown in Table S5. Since ΔF_{rot} depends on l, we find that $\Delta F_{\mathrm{rot}} \approx 0.8 k_{\mathrm{B}}T$ for the 120bp double helix, $\Delta F_{\mathrm{rot}} \approx 0.25 k_{\mathrm{B}}T$ for the bundle and 60bp helix and $\Delta F_{\mathrm{rot}} \approx 0.005 k_{\mathrm{B}}T$ for the tile and tetrahedron. This value may be compared with the $\Delta F_{\mathrm{trans}}$ contribution which is about $3 k_{\mathrm{B}}T$ in general, so that ΔF_{el} is about $1.65 k_{\mathrm{B}}T$ for the 120bp rod and $1 k_{\mathrm{B}}T$ for the tetrahedron under typical measurement conditions in this work. We emphasize that although the above analysis would appear to suggest requirement of knowledge of the unknown quantity q_{eff} at the outset, in practice this is not the case. This is because whereas ΔF_{el} depends linearly on q_{eff}, both $\Delta F_{\mathrm{trans}}$ and ΔF_{rot} depend weakly (logarithmically) on q_{eff}, and are therefore largely insensitive to q_{eff}[4]. Thus initial, approximate estimates of q_{eff} for each species suffice to determine $\Delta F_{\mathrm{trans}}$ and ΔF_{rot}, and the procedure may be repeated iteratively to refine the final estimate of measured effective charge q_{m}, if needed.

2.5: Calibration of the ET*e* measurement

We see from equation S4, that to make an accurate inference on q_{eff} from the inferred value of ΔF_{el} we also require accurate knowledge of ϕ_{m}, the electrical potential at the midplane of the slit. ϕ_{m} depends on the effective surface electrical potential at the silica walls of the slit, ϕ_{s} as follows

$$\phi_{\mathrm{m}} = 2\phi_{\mathrm{s}}\exp(-\kappa h) \tag{S11}$$

Treating ϕ_{s} as constant for experiments performed over a narrow range of salt concentrations (1.1-5.5 mM) in a given device, and combining equations S4 and S11 yields

$$\ln\frac{\Delta F_{\mathrm{el}}}{k_{\mathrm{B}}T} = \ln\frac{2q_{\mathrm{eff}}\phi_{\mathrm{s}}}{k_{\mathrm{B}}T} - \kappa h \tag{S12}$$

We measured ΔF_{el} for all the molecular species in our study under various conditions of salt concentration and slit height and plotted the measurements vs. a dimensionless system size, κh (Figure S7b). Fits of the data to equation S12 yield values for the y-intercept given by $2q_{\mathrm{eff}}\phi_{\mathrm{s}}$ in each case. Knowing the value of the surface potential ϕ_{s} in the device would permit us to determine q_{eff} for each molecular species. Since an accurate value of ϕ_{s} is not available *a priori* we perform a calibration measurement using a well-characterized molecular species, whose geometry is well defined and whose effective charge, q_{eff}, is known from calculation[14]. Here, we typically use a 60bp double-stranded DNA fragment carrying two fluorophore labels as a calibration molecule, for which we expect $q_{\mathrm{eff}} \cong -45\ e$ for measurements performed at ≈1 mM monovalent salt, similar to previous work[12,15]. Corresponding measured effective charge values, denoted as q_{m}, are displayed in Figure S7b. In general we find that the value of ϕ_{s} obtained for each device is stable over a timescale of several days and hence a single calibration run can be used for reliable measurements of effective charge, q_{m}, of further species under the same or similar conditions.

We point out that in order to enable a direct quantitative comparison of electrostatic interactions for the various molecules in the study, we work with measured effective charge values, q_{m}, rather than measured free energies, ΔF_{el}. This is because q_{m} is a property characteristic of the molecule and is robust to small variations in experimental conditions from one realization to the next. ΔF_{el} on the other hand, which is proportional to ϕ_{m}, is very sensitive to small variations in slit height and salt concentration, embodied in a single parameter κh as shown in equation S12 and discussed in Figure S7. Division of the measured

ΔF_{el} value by ϕ_{m} yields a measured value of effective charge, q_{m}, characteristic of the molecule of interest. Unlike t_{esc} and ΔF_{el}, q_{m} is robust to small changes in system size, κh, and permits direct comparison of measurements across independent realizations. Figure S7c displays individual measurements of ΔF_{el} converted to measured effective charge values, q_{m}, plotted vs. κh. Note that at $c = 1.5$ mM NaCl e.g., the slight dependence of η on salt concentration would imply $q_{\mathrm{eff}} \approx -46.2\, e$ for 60bp DNA, which is about 3% larger in magnitude than the value at 1 mM NaCl (outlined in section 3.3). However, η values for the DNA nanostructures are also expected to be similarly affected by salt concentration largely cancelling the slight influence of varying salt concentration in the final q_{eff} results (Figure S9). Measurements for all molecular species considered indeed show that q_{m} values remain constant within measurement uncertainty over a wider range of salt concentrations of ~1-6 mM (Fig. 1e).

2.6: Sources of error in q_{m}

In an ET*e* measurement that includes a calibration molecule, the uncertainty on t_{esc}, denoted as $t_{\mathrm{esc,e}}$, comes mainly from the statistical uncertainty of measuring *N* independent random events from a Poisson point process, and is given by $\frac{t_{\mathrm{esc}}}{\sqrt{N}}$ [11]. We point out that in an uncalibrated measurement a strong contribution to the measurement uncertainty stems from the uncertainty on κh, as described in detail previous work[4,12,14]. Use of a calibrator molecule in this work reduces the fractional measurement uncertainty and inaccuracy to a level essentially dictated by shot noise ($\frac{1}{\sqrt{N}}$).

As shown in previous work[12], the fractional error on q_{m}, may be estimated by

$$\frac{q_{\mathrm{m,e}}}{q_{\mathrm{m}}} = \frac{1}{\sqrt{N}} \frac{k_{\mathrm{B}}T}{\Delta F_{\mathrm{el}}} \tag{S13}$$

Since $\Delta F_{\mathrm{el}} \approx 1 - 2\ k_{\mathrm{B}}T$ in our measurements, we expect the fractional uncertainty on q_{m} to be comparable with the fractional statistical uncertainty on t_{esc}, which is simply $\frac{1}{\sqrt{N}}$. Measurements in this study are characterized by $N = 200 - 2000$ escape events, giving a measurement uncertainty of ≈2-7% on t_{esc} and on the molecular effective charge.

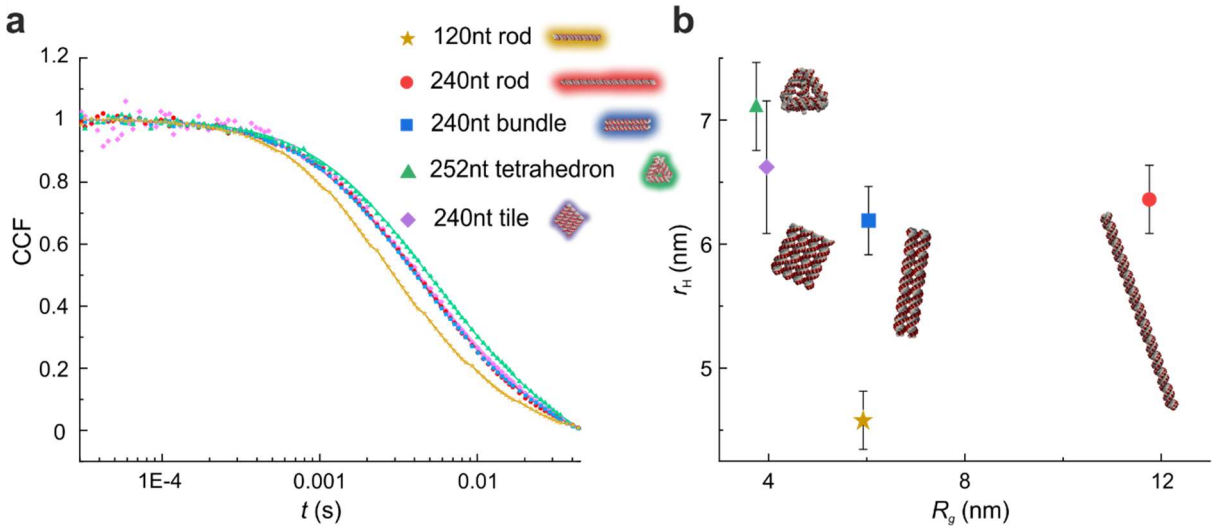


Figure S6. Measurement of hydrodynamic radii, r_H, by two-focus Fluorescence Correlation Spectroscopy (2fFCS). **(a)** The normalized cross-correlation function (CCF) was calculated for fluorescence emission recorded at separate detectors from labeled molecular species freely diffusing between two spatially separated, orthogonally polarized optical foci. A diffusion coefficient is extracted from fitting the CCF using the method outlined by Dertinger *et al.*[13], and the corresponding r_H value inferred from the Stokes-Einstein equation. (**b**) Measured r_H values for each nanostructure display little dependence on the molecular radius of gyration, R_g. Errors on r_H are evaluated as described in section 2.3.

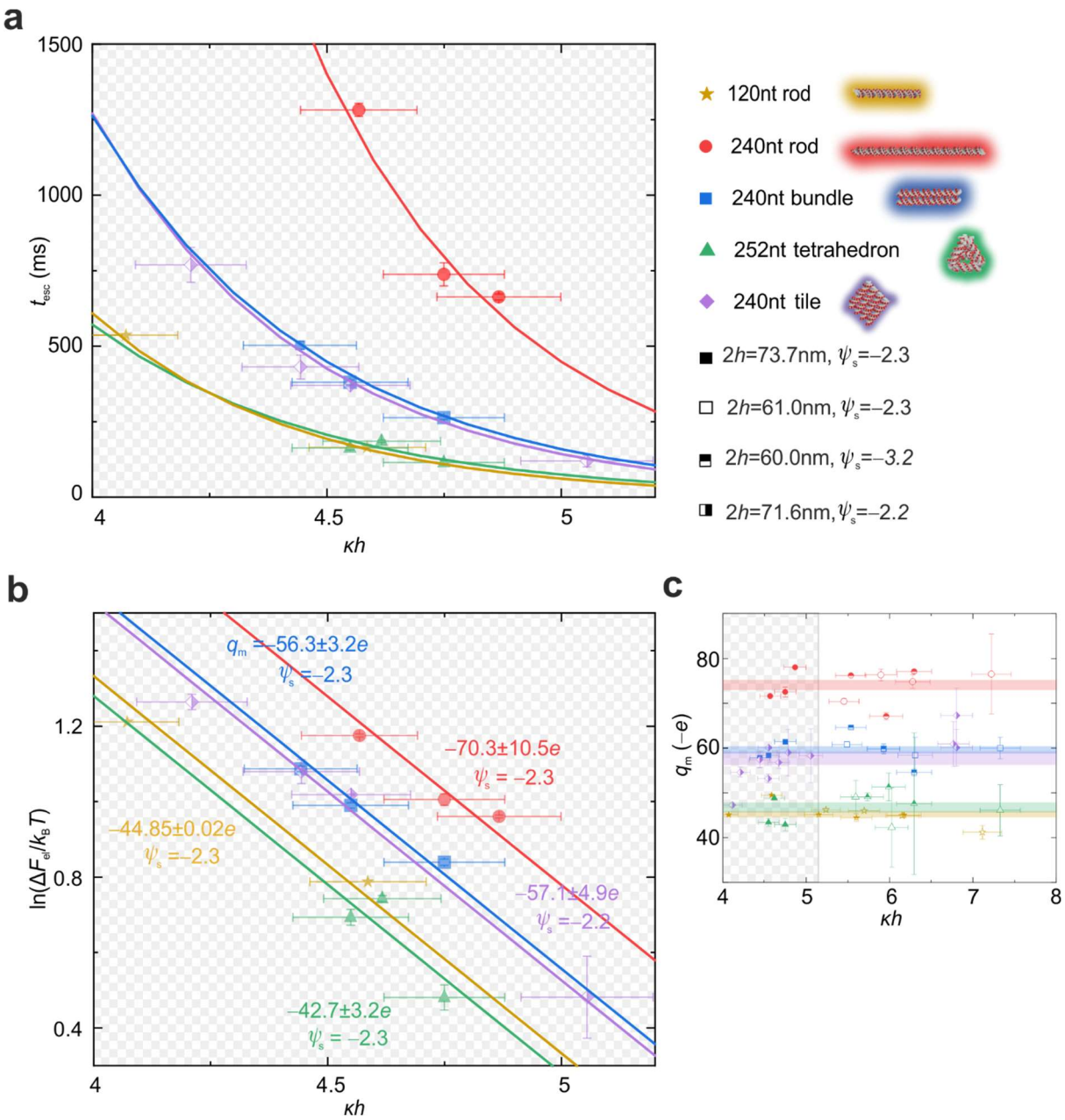


Figure S7. Measuring the effective charge of DNA nanostructures. (**a**) Escape times, t_{esc}, measured for the different DNA nanostructures in devices with slit depths $2h \approx$ 71.6 and 73.7 nm, and surface potentials $\psi_{\mathrm{s}} = -2.3$ to -2.2, over a range of Debye screening lengths, $\kappa^{-1} = 7 - 9$ nm, varied by changing the NaCl concentration, c. Error bars on t_{esc} values denote errors on fits of $P(\Delta t)$ vs. Δt data as shown in Figure 1 in the main text, while uncertainties in κh arise from accuracy limits on the measurement of slit height, h (uncertainty ~1 nm), discussed previously[9]. Lines are guides to eye. (**b**) Electrostatic trap depth (ΔF_{el}) inferred from the escape time measurements in (a), as described in the text. Fits of the data to equation S12 yield a measured molecular effective charge, q_{m}, with the silica surface potential, ψ_{s}, calibrated for each device using 60bp double-stranded DNA (as described in the text). (**c**) q_{m} values inferred for each individual measurement for all species are constant within measurement uncertainty for a given nanostructure species, over a range of salt concentrations ($c = 1.1 - 5.5$ mM) as well as across different measurement devices where slit heights ranged from $2h \approx 60.3$ to 73.7 nm and surface potentials ranged from $\psi_{\mathrm{s}} = -3.2$ to -2.2. Shaded region of the plot corresponds to the data range presented in (a) and (b). Data also shown in Figure 1e of the main text.

Procedure to calculate values of effective charge, q_{eff}, for DNA nanostructures

3.1: Free energy calculations to determine q_{eff}, for DNA nanostructures

As established in previous work[14,15] we determine the theoretical molecular effective charge, q_{eff}, by first calculating the electrostatic free energy difference, ΔF_{el}, associated with the trapping process. The electrostatic free energy, F, is determined from the dimensionless potential field, ψ, using the expression for the free energy functional

$$\frac{F}{k_{\mathrm{B}}T} = c\, N_{\mathrm{A}} \int_V [\kappa^{-2}|\nabla\psi|^2 - 2(\cosh\psi - \psi\sinh\psi - 1)]\, d\mathbf{x} \qquad \text{(S14)}$$

ψ is obtained by solving the non-linear Poisson-Boltzmann (NLPB) [equation S3] for the slit geometry (as shown in Figure 1a,c in the main text) using constant charge boundary conditions on all surfaces, including those of the slit and model DNA nanostructures, as described previously[14,15,18]. Here, the domain of interest is the electrolyte denoted by V, c is the bulk concentration of ionic species, and N_{A}, Avogadro's number. The integral in equation S14 is evaluated for the two trapping states: (1) Particle at the midplane of the slit region, oriented with the longest axis along the midplane ($\theta = 90°$), which represents the minimum electrical energy state of the object in the slit and (2) Particle at the minimum energy location in the "pocket" region. The difference between the two values yields the quantity that is measurable in the context of ET*e* experiments, i.e. $\Delta F_{\mathrm{el}} = F_1 - F_2$ (Figure S8). We have previously shown that ΔF_{el} is related to the effective charge via the midplane electrical potential in the slit, i.e., $q_{\mathrm{eff}} = \Delta F_{\mathrm{el}}/\phi_{\mathrm{m}}$[14].

Note that for calculations of the electrical free energy of a charged spheroid or the square-tile DNA nanostructure in bulk solution, as referred to in the analyses in Figures 3 and 4 of the main text, we evaluate equation S14 for a single state of the object immersed in an electrolyte. Here we solve the NLPB equation for the object immersed in a rectangular domain consisting only of bulk electrolyte, whose bounding surfaces are situated at a minimum

separation of $\approx 5\kappa^{-1}$ from the charge object, as shown in Figure S8. A zero electric field boundary condition on the domain surfaces ensures electroneutrality within the system.

3.2: Electrostatic modeling of DNA Nanostructures: molecular geometry and boundary conditions

We used COMSOL Multiphysics to solve equation S3 and evaluate equation S14 for DNA nanostructures in our device geometry. We constructed prototypes of origami structures from assemblies of cylinders, each representing a segment of double-helix in the nanostructure. We assume that the total structural charge, q_{str}, of the molecule is uniformly distributed over the surface area of the cylinders constituting the origami models, A_{m}; i.e. $\sigma_{\mathrm{m}} = q_{\mathrm{str}}/A_{\mathrm{m}}$. The radii of the cylinders, r_{cyl}, were assumed to be 12 Å, and the length of each cylinder, l, was determined by nb, where n is the number of base pairs in the particular segment of double-helix and $b = 3.4$ Å is the rise per basepair for a B-DNA double-helix. Thus a uniform, constant surface charge density, σ_{m}, was used as a boundary condition for the nanostructures. For the bundle and square-tile nanostructures, the nominal centre to centre separation between cylinders was assumed to be $s \approx 2.7$ Å[3]. However for the square-tile, we examined the influence of this parameter on the calculated value of ΔF_{el} by varying s over a range of approximately 2.6–4.5 Å (see main text Figure 4).

We neglect angular and spatial degrees of freedom in the calculation of ΔF_{el} (N.B. these degrees of freedom are removed from the experimental measurement prior to comparison with the theoretical value of ΔF_{el}, as described in section 2.4). We thus assume that the centre of mass of each DNA nanostructure lies in the midplane (local minimum of the electrostatic free energy) of a slit of height $2h = 75$ Å. For all planar nanostructure configurations (cylinder, bundle and square) we further assumed that the axes of all cylinders lie on the midplane of the slit (Figure S8b). However, in the special case of tetrahedron, the minimum energy configuration had one of the faces of tetrahedron parallel to the slit walls, with the centre of

geometry of the tetrahedron contained in the midplane of the slit (Figure S8c). Figure 2 of the main text presents these models together with electrical potential distributions both on the DNA nanostructure surfaces as well as in the electrolyte surrounding the model nanostructures in the pocket region. The boundary condition for the silica slit walls was a constant surface charge density, $\sigma_{\mathrm{w}} = -0.1e$ nm^{-2}, which is a typical value for silica interfaces in contact with aqueous electrolytes under our experimental conditions, as noted in previous work. We emphasize that the value of σ_{w} does not affect the results for q_{eff}.

A similar procedure was applied to the charged spheroids presented in the main text. For all spheroids, the centre of mass and major axis of the objects lie on the midplane of the slit. The aspect ratio of the spheroids was varied by changing the values of major and minor axes while constraining the volume of all objects to the volume of a sphere of radius 4 nm, corresponding to the radius of gyration of the DNA tetrahedron. A total structural charge of $q_{\mathrm{str}} = -240e$ was distributed uniformly over the surface of the spheroids, corresponding to a constant uniform charge density boundary condition on the particle surface.

3.3: Effect of salt concentration on effective charge s and electrostatic modelling of DNA Nanostructures

Figure S9 presents the impact of salt concentration on the calculated effective charge, q_{eff}, for the various DNA nanostructures. Taking $c = 1.5$ mM as a reference ionic strength, we normalize q_{eff} calculated at a salt concentration c for each DNA nanostructure to its corresponding value at $c = 1.5$ mM which we denote as q_0. A 30-50% change in salt concentration, c (corresponding to a change of ± 0.5 mM) alters the effective charge for 60bp double-stranded DNA by about $\pm 3\%$ which is of the same order or smaller than our measurement uncertainty on q_{m}. The renormalization factor, η, can be expected to vary substantially with salt concentration, e.g., by a factor of up to 4 over a change of 4 orders of magnitude in salt concentration for infinitely long rigid rods[14,19,20]. However, the range of

variation in η over our experimentally probed range of concentrations is of the order of 10% as shown in Fig. S9. Furthermore, we find a similar qualitative dependence of q_{eff} on salt concentration for all DNA nanostructures, which suggests that a large portion of the small salt concentration dependence of q_{eff} cancels out when the nanostructure measurements are referenced to that of the 60bp DNA calibrator molecule at each salt concentration

3.4: Comparison of renormalization factors calculated within our theoretical model with prior theoretical estimates

We have compared our calculated values of η to the results for infinite charged rods from previous theoretical work[21] and have found that our value of $\eta = 0.34$ for the 240nt case agrees quite well with the published value for infinite rods $\eta \approx 0.3$. Here the length of the rod is significantly longer than the Debye length and end effects could be expected to be less important. For the 240nt bundle, however the length of the effective cylinder is not substantially larger than the Debye length and end effects are expected to matter[21]. We find that for the bundle, our calculated of $\eta = 0.25$ agrees very well with our measurements, but is indeed slightly different from the value $\eta \approx 0.19$ suggested by Téllez & Trizac[22]. Here we use an assumption of equal surface areas in order to map the geometry of the bundle on to that of a simple cylinder. Finally, for the 120nt rod, our model of effective charge suggests $\eta = 0.38$, where the infinite rod model suggests $\eta \approx 0.3$. Overall, the comparison of our calculations for finite-length rods with a calculation for an infinite rod is not expected to be favourable for smaller rod lengths since end effects do begin to matter as the rod length approaches the Debye length.

We further note that the results for the square-tile nanostructure may be compared to those obtained for the idealized "thin disk". In order to enable such a comparison, we mapped the geometry of our square-tile on to that of a disc whose surface area corresponds to that of the square-tile and estimated an effective radius a for the square-tile. In the regime where $\kappa a <$

1, our calculated value of $\eta = 0.22$ can be compared with $\eta = 0.18$ obtained by Agra et al.[23]. Note that exact agreement is not expected given the approximations involved in making such a comparison.

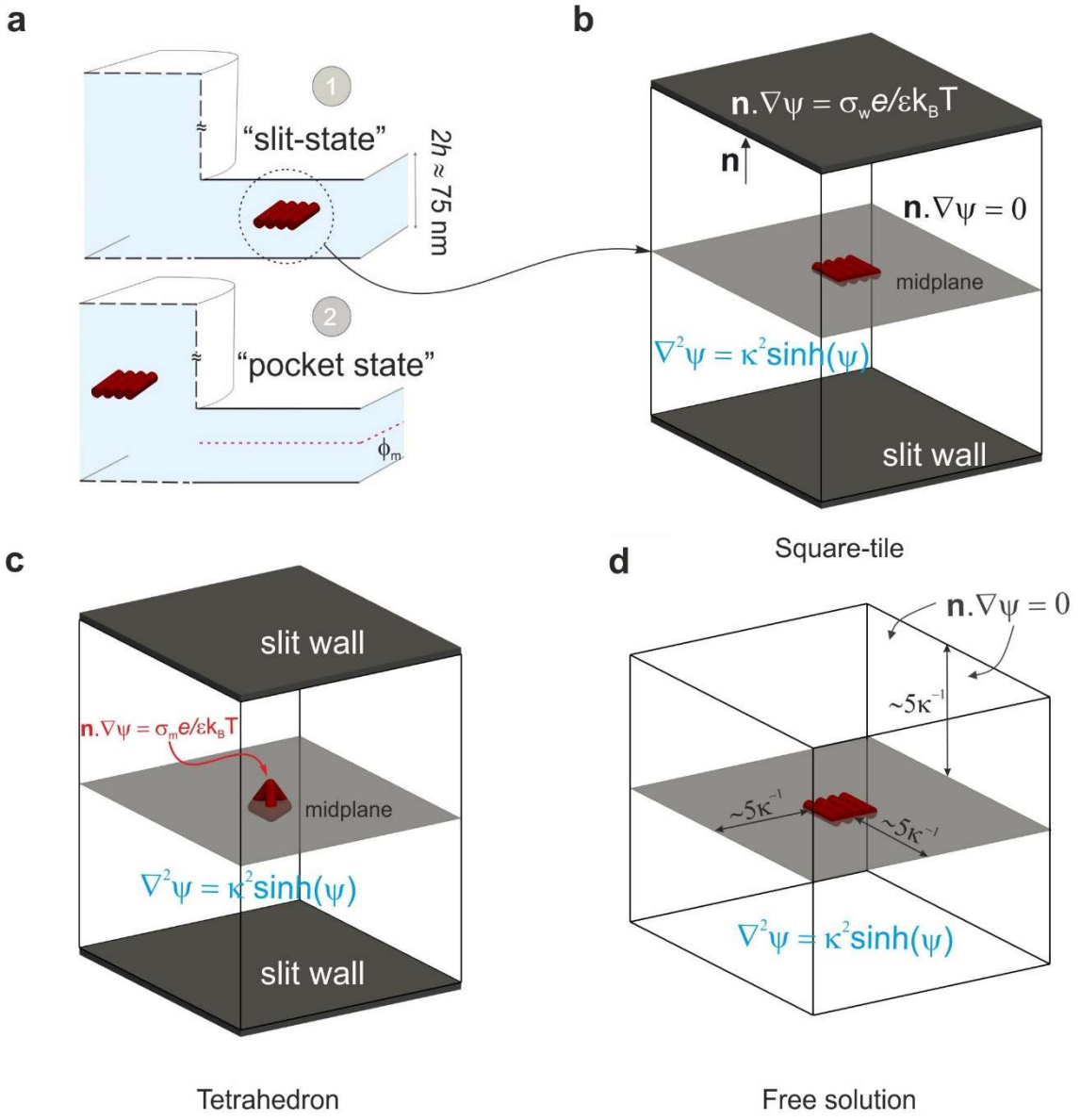


Figure S8. Schematic representations of DNA nanostructures in PB free energy calculations. (**a**) State (1) shows a square-tile origami in the slit region, and state (2) represents the trapped object in the pocket. (**b**) Position and orientation of a square-tile nanostructure in the slit state for free energy calculations. The governing PB equation in the electrolyte and the constant surface charge boundary conditions applied to the DNA nanostructures and slit surfaces are depicted. Note that n refers to the unit surface normal pointing out of the electrolyte. (**c**) Placement and orientation of a tetrahedron nanostructure in the slit for free energy calculations. (**d**) Geometry used for calculations of electrostatic free energy, *F*, in bulk solution for spheroids and square-tile structures shown in the main text Figures 3 and 4.

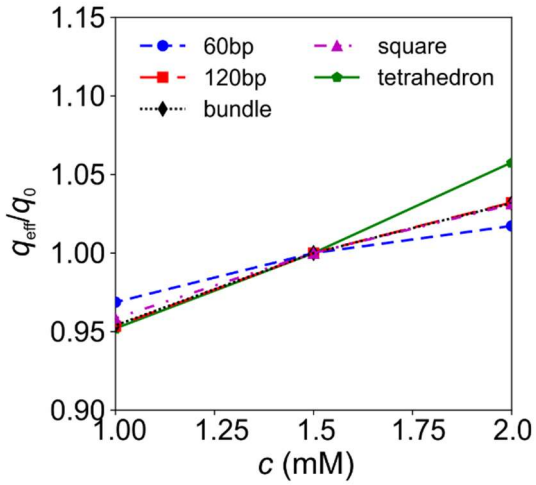


Figure S9. Impact of salt concentration on the effective charge, q_{eff}. Variation in salt concentration around a reference value ($c = 1.5$ mM) leads to small changes in the calculated effective charge. Graphs for the various DNA nanostructures display the ratio $\eta = q_{\mathrm{eff}}/q_0$ as a function of salt concentration, c, where q_0 is the effective charge of each nanostructure species at a reference salt concentration of $c = 1.5$ mM. $c = 1 - 2$ mM corresponds to $\kappa h = 3.7 - 5.2$, similar to the experiment shown in Figure S8a,b. For $c = 1.5$ mM we have $q_{\mathrm{eff}} = -46.2\ e$ for 60dsDNA.

REFERENCES

(1) Green, M. R.; Sambrook, J. Precipitation of DNA with ethanol. *Cold Spring Harbor Protocols* **2016**, *12*, 1116-1120.

(2) Goodman, R. P.; Schaap, I. A. T.; Tardin, C. F.; Erben, C. M.; Berry, R. M.; Schmidt, C. F.; Turberfield, A. J. Rapid chiral assembly of rigid DNA building blocks for molecular nanofabrication. *Science* **2005**, *310* (5754), 1661-1665.

(3) Fischer, S.; Hartl, C.; Frank, K.; Radler, J. O.; Liedl, T.; Nickel, B. Shape and interhelical spacing of DNA origami nanostructures studied by small-angle x-ray scattering. *Nano Lett.* **2016**, *16* (7), 4282-4287.

(4) Ruggeri, F.; Krishnan, M. Entropic trapping of a singly charged molecule in solution. *Nano Lett.* **2018**, *18* (6), 3773-3779.

(5) Doty, D.; Lee, B. L.; Stérin, T. scadnano: a browser-based, scriptable tool for designing DNA nanostructures. **2020**, 11841, *arXiv*. https://arxiv.org/abs/2005.11841 (accessed October 14, 2020).

(6) Ouldridge, T. E.; Louis, A. A.; Doye, J. P. K. Structural, mechanical, and thermodynamic properties of a coarse-grained DNA model. *J. Chem. Phys.* **2011**, *134* (8), 085101.

(7) Poppleton, E.; Bohlin, J.; Matthies, M.; Sharma, S.; Zhang, F.; Šulc, P. Design, optimization and analysis of large DNA and RNA nanostructures through interactive visualization, editing and molecular simulation. **2020**, *48* (12), e72.

(8) Poppleton, E.; Romero, R.; Mallya, A.; Rovigatti, L.; Sulc, P. OxDNA.org: a public webserver for coarse-grained simulations of DNA and RNA nanostructures. *Nucleic Acids Research* **2021**, *49* (W1), W491-W498.

(9) Ruggeri, F.; Zosel, F.; Mutter, N.; Różycka, M.; Wojtas, M.; Ożyhar, A.; Schuler, B.; Krishnan, M. Single-molecule electrometry. *Nature Nanotechnology* **2017**, *12* (5), 488-495.

(10) Krishnan, M.; Mojarad, N.; Kukura, P.; Sandoghdar, V. Geometry-induced electrostatic trapping of nanometric objects in a fluid. *Nature* **2010**, *467* (7316), 692-695.

(11) Ruggeri, F.; Krishnan, M. Spectrally resolved single-molecule electrometry. *J. Chem. Phys.* **2017**, *148* (12), 123307.

(12) Ruggeri, F.; Krishnan, M. Lattice diffusion of a single molecule in solution. *Physical Review E* **2017**, *96* (6), 062406.

(13) Dertinger, T.; Pacheco, V.; von der Hocht, I.; Hartmann, R.; Gregor, I.; Enderlein, J. Two-focus fluorescence correlation spectroscopy: a new tool for accurate and absolute diffusion measurements. *ChemPhysChem* **2007**, *8* (3), 433-443.

(14) Krishnan, M. A simple model for electrical charge in globular macromolecules and linear polyelectrolytes in solution. *J. Chem. Phys.* **2017**, *146* (20), 205101.

(15) Krishnan, M. Electrostatic free energy for a confined nanoscale object in a fluid. *J. Chem. Phys.* **2013**, *138* (11), 114906.

(16) Kjellander, R. & Mitchell, D. J. Dressed ion theory for electric double layer structure and interactions; an exact analysis. *Mol. Phys*. **1997**, *91* (2), 173-188

(17) Ulander, J., Greberg, H. & Kjellander, R. Primary and secondary effective charges for electrical double layer systems with asymmetric electrolytes. *J. Chem. Phys.* **2001**, *115* (15), 7144-7160.

(18) Theodoor, J.; Overbeek, G. The role of energy and entropy in the electrical double layer. *Colloids and Surfaces* **1990**, *51*, 61-75.

(19) Netz, R. R. & Orland, H. Variational charge renormalization in charged systems. *Eur. Phys. J. E Soft Matter* **2003**, *11* (3), 301-311.

(20) Aubouy, M., Trizac, E. & Bocquet, L. Effective charge versus bare charge: an analytical estimate for colloids in the infinite dilution limit. *J. Phys. A* **2003**, *36* (22), 5835-5840.

(21) Manning, G. S.; Mohanty, U. Counterion condensation on ionic oligomers. *Physica A* **1997**, *247* (1), 196-204.

(22) Téllez, G.; Trizac, E. Exact asymptotic expansions for the cylindrical Poisson–Boltzmann equation. *J. Stat. Mech.: Theory Exp.* **2006**, *2006* (06), P06018-P06018.

(23) Agra, R.; Trizac, E.; Bocquet, L. The interplay between screening properties and colloid anisotropy: Towards a reliable pair potential for disc-like charged particles. *Eur. Phys. J E* **2004**, *15* (4), 345-357.

Appendix H

Supplementary materials for "Single-molecule size-and-shape spectrometry on a chip"

Supplementary Materials for

Single-molecule size-and-shape spectrometry on a chip

Authors: Xin Zhu[1]†, Timothy J. D. Bennett[1]†, Konstantin C. Zouboulis[1,2], Dimitrios Soulias[1], Michal Grzybek[3], Justin L. P. Benesch[1,2], Afaf H. El-Sagheer[4,5], Ünal Coskun[3,6-8], Madhavi Krishnan[1,2]*

Affiliations:

[1] Physical and Theoretical Chemistry Laboratory, Department of Chemistry, University of Oxford; South Parks Road, Oxford OX1 3QZ, United Kingdom

[2] The Kavli Institute for Nanoscience Discovery, University of Oxford; Sherrington Road, Oxford OX1 3QU, United Kingdom

[3] Center of Membrane Biochemistry and Lipid Research, University Hospital and Faculty of Medicine Carl Gustav Carus, Technical University Dresden; 01307 Dresden, Germany

[4]School of Chemistry and Chemical Engineering, University of Southampton; Highfield, Southampton, SO17 1BJ, United Kingdom

[5]Institute for Life Sciences, University of Southampton; Highfield Campus, Southampton, SO17 1BJ, United Kingdom

[6]Paul Langerhans Institute Dresden (PLID) of the Helmholtz Center Munich at the University Hospital Carl Gustav Carus and Faculty of Medicine, Technical University Dresden; 01307 Dresden, Germany

[7]German Center for Diabetes Research (DZD), 85764 Neuherberg, Germany

[8]Max Planck Institute of Molecular Cell Biology and Genetics, 01307 Dresden, Germany

*Corresponding author. Email: madhavi.krishnan@chem.ox.ac.uk

† These authors contributed equally to this work

The PDF file includes:

Materials and Methods
Supplementary Text
Figs. S1 to S22
Tables S1 to S5
References 73-96

Other Supplementary Materials for this manuscript include the following:

Movies S1 to S2

Table of Contents

Materials and Methods 5
S1.1. Commercially obtained samples 5
Protein labelling procedure 5
Electrospray mass spectrometry (ESI-TOF) 5
Native mass spectrometry method 6
S1.2. DNA nanostructure synthesis 6
S1.3. SAM-IV riboswitch and reference DNA synthesis 6
Selective β-cyanoethyl Removal 6
DNA Deprotection 7
RNA Deprotection (2'-O-TC) 7
RP-HPLC Purification 7
Post-synthetic oligonucleotide modification 7
Oligonucleotide mass spectrometry 7
S1.4. Insulin receptor Ectodomain (IR-ECD) purification and labelling 7
Supplementary Text 8
S2. ETsy experimental procedure 8
S2.1. Fluorescence Microscopy 8
S2.2. Data analysis to determine the escape time of a molecular species 8
Fitting histograms of escape-time data 8
Determining number of timescale components characterizing an escape-time histogram 9
S2.3. Nanoslit surface passivation procedure 10
S2.4. High-throughput single-molecule counting down to femtomolar concentrations 10
S3. Theoretical model relating average escape time to molecular properties r_H and D_s 11
S4. ETsy measurements of biomolecules 12
S4.1. Proteins 12
S4.2. Dye derivatives 13
S4.3. DNA 13
S5. Independent estimates of biomolecular properties from 3D molecular models and FCS 14
S5.1. Stokes' radius r_H using HYDROPRO 14
S5.2. Measured Stokes' radius $r_{H,FCS}$ using Dual-focus fluorescence correlation spectroscopy (2f-FCS) 14
S5.3. Theoretical estimates of r_H for the nucleic acid double helix 15
S5.4. Determination of r_H and D_s relation for the SAM-IV riboswitch 15
S5.5. Determining D_s from molecular 3D structures 16

S6. ETsy measurements of biomolecular structure 16
S6.1. BD Simulations on spherical particles 17
S6.2. Calibration of the measurement using globular proteins 17
S6.3. Structural measurements on double-stranded nucleic acids 18
S6.4. Structural measurements on DNA nanostructures 18
S6.5. Structural inferences on the SAM-IV riboswitch 19
S6.6. Structural measurements on IR-ECD 19
S7. Single-molecule ETsy 20
S7.1. Single-particle tracking and analysis 20
S7.2. Validation of single molecule spectrum approach using dsDNA 21
S7.3. Measurements on the SAM-IV riboswitch 21
S8. Assessing the extent of molecular binding in a bi-molecular reaction 22
S8.1. Determining the fractional abundance of the bound-state, m_2 22
S8.2. Inferring the equilibrium constant, K_d 23
S8.3. Measurements of on- and off-rate constants in a binding reaction 23
S8.4. Estimating the overall uncertainty in measurements of m_2 24
S8.5. DNA hybridisation 24
Equilibrium measurements of affinity constants 25
Measurements of reaction kinetics 25
Calculation of theoretically expected binding affinities 26
S8.6. 24bp + EcoRI binding affinity measurement 26
S8.7. HLA + IgG binding affinity measurement 26
S8.8. Insulin aptamer + Insulin binding affinity measurement 27
S8.9. Insulin + IR-ECD binding affinity measurement 27
S8.10. IR-ECD + Insulin binding affinity measurement 27
S8.11. Sources of inaccuracy in molecular binding measurements 28
S9. Supplementary Tables 29
Table S1 Nucleic acid sequences 29
Table S2 Disordered Protein sequences 31
Table S3 Details of protein labelling 32
Table S4 Measured and structural estimates of Stokes' radii of proteins 33
Table S5 Measured Stokes' radii of fluorescently labelled nucleic acids using 2f-FCS 34
S10. References 57

List of Figures

Fig. S1 LC-MS spectra of ATTO532 dye derivatives.

Fig. S2 Intact mass spectrometry analysis of ATTO532-labelled insulin.

Fig. S3 Native mass spectrometry analysis of ATTO532-labelled HLA-A*03:01.

Fig. S4 LC-MS spectra of (A) 119 nt reference DNA and (B) SAM-IV riboswitch.

Fig. S5 Passivation of slits with m-PEG-silane and demonstration of low copy-number molecular counting.

Fig. S6 Multiexponential histogram fitting and ground-truth validation.

Fig. S7 BD simulations of spherical particles diffusing in a nanostructured landscape.

Fig. S8 Measurements on Apoferritin (FER), disordered proteins and ATTO542 dye derivatives.

Fig. S9 Comparison of Stokes' radii measured using ETsy and 2f-FCS for a set of globular proteins (Fig. 2D, main text).

Fig. S10 Measurements of the rise-per basepair, b, in dsDNA and dsRNA.

Fig. S11 Measurement of D_s and r_H for DNA nanostructures.

Fig. S12 Measurement of the structural properties of the SAM-IV riboswitch.

Fig. S13 Measurement of D_s and r_H of the IR-ECD.

Fig. S14 Validation of single-molecule spectra measurements using dsDNA.

Fig. S15 Single-molecule tracking of the SAM-IV riboswitch.

Fig. S16 Single-molecule measurement of the SAM-IV riboswitch and binding of S-adenosyl methionine (SAM) to the riboswitch.

Fig. S17 Assessment of the attainment of equilibrium in the measurement of K_d for a 15 base ssDNA oligo binding to 200 base ssDNA.

Fig. S18 Measurement of on and off-rates and assessment of the role of cooperative unbinding in k_{off} measurements.

Fig. S19 Comparison of K_d measurements using bi-exponential and t_{av} fitting methods to extract m_2.

Fig. S20 Measuring K_d for the interactions of EcoRI, HLA-IgG and IR-ECD and their binding partners.

Fig. S21 Measuring K_d for the binding of insulin to an Insulin aptamer.

Fig. S22 Measuring K_d for the interaction of IR-ECD with insulin.

Materials and Methods

S1.1. Commercially obtained samples

Recombinant human insulin (INS, 91077C) was obtained from SAFC Biosciences, Andover, UK. Ubiquitin from bovine erythrocytes (Ub, U6253), Thioredoxin from Escherichia coli (TRX, T0910), Ribonuclease A from bovine pancreas (RNase A, R6513), Myoglobin from equine heart (MB, M1882), β-Lactoglobulin A from bovine milk (LGB, L7880), Carbonic anhydrase (CA, C7025), human apo-Transferrin (TF, T2036) and Apoferritin (FER, A3660) were all from Sigma-Aldrich, Gillingham, UK. Human insulin detector antibody (IgG, ab253508) was purchased from Abcam, Cambridge, UK. HLA-ABC monoclonal antibody W6/32 (IgG, 16-9983-85) and Human EGF Recombinant Protein (PHG0314) were from Thermo Fisher Scientific. HLA-A*03:01 RLRAEAQVK was provided by the NIH Tetramer Core Facility at Emory University, Georgia, USA. EcoR1 (R0101M) was obtained from New England Biolabs. Nucleic Acid samples were mostly obtained from Integrated DNA Technologies Inc., Coralville, Iowa, with sequences shown in Table S1. With the exception of the SAM-IV riboswitch sample, and reference DNA sample – synthesis of these is outlined in section S1.3. All fluorescent dyes were obtained as lyophilized powder from ATTO-TEC GmbH, Siegen, Germany. The disordered proteins Starmaker-like (Stm-l), Prothymosin α (ProTα), were obtained as described in Ref (*20*). A 13-mer proline-rich polypeptide was obtained by custom synthesis from Bio-Synthesis Inc., respectively. Details of the sequences of these proteins can be found in Table S2. Samples were generally prepared using UltraPure DI water (Invitrogen) and PBS (pH 7.4, Gibco). Samples, except for antibodies were aliquoted, flash-frozen and stored at either -20 or -80°C as per manufacturer's instructions.

Protein labelling procedure

To protein stock solutions in DI water or PBS (pH 7.4), 1 M $NaHCO_3$ (Sigma-Aldrich) was added to a final concentration of 0.1 M in order to adjust the pH to 7.5 - 8.3. After addition of a molar excess of ATTO532 NHS-ester in DMSO or DI water, the reaction was incubated for 1 hour at room temperature, protected from light. Unreacted dye was removed by buffer exchanging twice into PBS (pH 7.4) using Zeba micro spin desalting columns with 7 K or 40 K MWCO (Thermo Fisher Scientific) following the manufacturer's instructions. Labelled insulin and dye were separated into PBS using size exclusion chromatography on a Superdex 75 10/300 GL column (Cytiva) on an AKTA pure protein purification system (Cytiva). Post-labelling concentrations and degree-of-labelling (DOL) were calculated from absorbance measurements 280 nm and 532 nm on a Nanodrop spectrophotometer (Thermo Fisher Scientific), and are presented in Table S3.

Electrospray mass spectrometry (ESI-TOF)

Reverse-phase chromatography on an Agilent 1100 HPLC system (Agilent Technologies, Palo Alto, CA, USA) was coupled in-line to a 1969 MSD-ToF electrospray ionisation orthogonal time-of-flight mass spectrometer (Agilent Technologies Inc, Palo Alto, CA, USA). Samples were diluted to a concentration of 0.02 mg/ml (for insulin) or 10 µM (for ATTO532 dyes) in 0.1% formic acid. 50 µL were injected on to a 2.1 mm x 12.5 mm *Zorbax* 5 µm 300SB-C3 guard column. The column oven was maintained at 40^{o}C. Ultra-high purity water (Optima, LC-MS grade, Fisher Chemical) with 0.1% formic acid (Optima, LC/MS grade, Fisher Chemical) (solvent A) and methanol (Optima, LC/MS grade, Fisher Chemical) with 0.1% formic acid (solvent B) were used as mobile phases. Chromatography was started at 90% A and 10% B for 15 s, followed by a two-stage linear gradient from 10% B to 80% B over 45 s and from 80% B to 95% B over 3 s. Elution was performed isocratically at 95% B for 1 min 12 s followed by

equilibration at 90% A and 10% B over 45 s. The flow rate was set to 1.0 mL/min. For mass spectrometry the instrument was configured with the standard ESI source and operated in positive ion mode. The capillary voltage was set to 4000 V and the nebulizer pressure to 60 psi. The source temperature was kept at 350°C with a drying gas flow rate of 12 L/min. The ion optic voltages were: fragmentor 250 V, skimmer 60 V and octopole RF 250 V. Data analysis was performed using Agilent MassHunter Qualitative Analysis B.07.00 software (Agilent Technologies, Palo Alto, CA, USA), see Fig. S1 and Fig. S2.

Native mass spectrometry method

HLA was desalted and buffer exchanged twice into 250 mM ammonium acetate using Zeba micro spin desalting columns with 7 K MWCO following the manufacturer's instructions and diluted to 10 µM. Spectra were acquired on a Q Exactive UHMR Orbitrap mass spectrometer (Thermo Fisher Scientific) with in-house prepared gold-plated capillaries using the following parameters: capillary voltage of 1.3 kV, capillary temperature of 100°C, in-source trapping of -10 V and HCD energy 1. Spectra were averaged and deconvolved using UniDec software (*73*), see Fig. S3.

S1.2. DNA nanostructure synthesis

Two DNA nanostructures were synthesized as described in Ref. (*36*). The "square tile" consists of four terminally tethered 30 base-pair double-stranded (ds) DNA helices, and the "bundle" consists of two 60ds DNA helices tethered at two points along the length of the helix. Both structures consist of 240 nucleotides (nt), and are labelled with a single ATTO532 dye.

S1.3. SAM-IV riboswitch and reference DNA synthesis

Single stranded (ss) RNA and DNA sequences, both 119 nt in length, were synthesized (sequence information in Table S1) on an Applied Biosystems 394 automated DNA/RNA synthesizer using a standard phosphoramidite cycle of detritylation, coupling, capping, and oxidation on a 0.2 µmole scale. For DNA, TCA (3% in dichloromethane), 1*H*-tetrazole (0.45 M in acetonitrile), Cap A (10% acetic anhydride, 10% lutidine and 80% tetrahydrofuran) / Cap B (16% N-methylimidazole in tetrahydrofuran) and iodine (0.02 M in tetrahydrofuran, pyridine and water) were used. Pre-packed nucleoside SynBaseTM CPG 3000/110 (Link Technologies) resins were used and β-cyanoethyl protected phosphoramidites (dA-bz, dG-ib, dC-bz and dT where bz = benzoyl and ib = *iso*-butyryl, Sigma-Aldrich) were dissolved in anhydrous acetonitrile (0.1 M) immediately prior to use. The coupling time for dA, dC, dG and dT monomers was 45 s, and 600 s for the 5'-amino modifier phosphoramidite monomer. Stepwise coupling efficiencies were determined by automated trityl cation conductivity monitoring and were >98% in all cases.

For RNA, 2'-O-TC protected RNA phosphoramidites (A-bz, C-ac, G-ib, U, Sigma-Aldrich) monomers were used. Monomers were dissolved in anhydrous toluene: acetonitrile (1:1 v/v) (0.1 M) immediately prior to use. Capping and oxidation reagents were identical to those used in DNA synthesis while ethylthiotetrazole (ETT) (0.25 M in acetonitrile, Link Technologies) was used as a coupling reagent. The coupling time for all monomers during RNA synthesis was 3 min. Stepwise coupling efficiencies were determined by automated trityl cation conductivity monitoring and in all cases were >97%.

Selective β-cyanoethyl Removal

RNA and DNA bearing the 5'-primary amines were treated on-column with diethylamine (20% in anhydrous acetonitrile) for 20 min at room temperature. The resin was then washed with acetonitrile (3 x 1 mL) and dried with argon.

DNA Deprotection

DNA was cleaved from solid support and deprotected by exposure to a concentrated solution of aqueous ammonia in a sealed vial for 5 h at 55 °C. After drying *in vacuo*, oligonucleotides were dissolved in water and subject to further purification.

RNA Deprotection (2'-O-TC)

The solid support was exposed to dry ethylenediamine: toluene (1:1 v/v) for 6 h at room temperature, washed with toluene (3 x 1 mL) then acetonitrile (3 x 1 mL), and dried using argon. The cleaved RNA was eluted with water and purified using RP-HPLC.

RP-HPLC Purification

DNA and RNA and were purified using an Agilent system with Kinetex® C18 column (10 mm x 250 mm, pore size 100 Å, particle size 5 μm), a gradient of buffer A (0.1 M TEAA, pH 7.5) to buffer B (0.1 M TEAA, pH 7.5 50% v/v acetonitrile), and flow rate of 5 mL/min. A gradient of 0%-50% buffer B over 28 min at 55°C was used in all cases. Following HPLC, Amicon®Ultra-4 0.5 mL centrifugal filters (Merck, cat. no. UFC5x) were used to desalt and concentrate oligonucleotide samples.

Post-synthetic oligonucleotide modification

Freeze-dried oligonucleotide (20 nmol) was dissolved in $NaHCO_3$ buffer (0.5 M, pH 8.5, 30 μL) and mixed with ATTO NHS ester (200 nmol, 20 μL) dissolved in DMSO. The reaction was then left for 4 h at 25°C with 750 rpm shaking. After dilution with water, the samples were desalted using NAP™-10 Columns (Cytivia, cat. no. 17085402) according to the manufacturer's instructions. RNA or DNA oligonucleotides were purified using HPLC as described above.

Oligonucleotide mass spectrometry

All DNA and RNA were characterized by negative-mode electrospray using a UPLC-MS Waters XEVO G2-QTOF mass spectrometer and an Acquity UPLC system with a BEH C18 1.7 μm column (Waters). A gradient of methanol in triethylamine (TEA) and hexafluoroisopropanol (HFIP) was used (buffer A, 8.6 mM TEA, 200 mM HFIP in 5% methanol/water (v/v); buffer B, 20% v/v buffer A in methanol). Buffer B was increased from 0–70% over 7.5 min or 15–30% over 12.5 min for normal oligonucleotides and 50–100% over 7.5 min for hydrophobic oligonucleotides. The flow rate was set to 0.2 mL/min. Raw data were processed and deconvoluted using the deconvolution software MassLynx v4.1 and Unidec (*73*), see Fig. S4.

S1.4. Insulin receptor Ectodomain (IR-ECD) purification and labelling

The Insulin Receptor ectodomain (IR-ECD) was purified as previously described by Gutmann et al. (*61*). For labeling fivefold molar excess of Alexa Fluor 532 NHS ester (A20001, Thermo Fisher Scientific) was used. The reaction was performed on ice for 1 h and was quenched with a fivefold molar excess of buffered ethanolamine over Alexa Fluor 532 NHS. The labelled protein was separated from free dye on a Superdex 200 Increase 10/300 GL column equilibrated with 25 mM Hepes, 150 mM NaCl, pH 7.4. The protein was aliquoted and snap-frozen with liquid nitrogen.

Supplementary Text

S2. ETsy experimental procedure

S2.1. Fluorescence Microscopy

Microfluidic devices (Fig. S5A) were fabricated using procedures similar to those outlined in previous studies (*12*, *20*, *39*, *74*), with the exception that in this study the silicon and glass substrates were thermally rather than anodically bonded. 10-100 µL of sample is loaded into one pair of loading reservoirs (Fig. S5B), while the other pair of reservoirs are sealed and connected to a vacuum pump. The pressure drop across the length of the channels loads the molecules into the trapping area under steady Poiseuille flow. Movies of fluorescent molecules diffusing through the trapping nanostructures are recorded either with the pressure drop applied (under flow), or without, depending on the modality of experiment. Movies are recorded for 1 to 10 minutes using an epi-fluorescent wide-field fluorescent microscope at a frame rate of 100 Hz, generally, with an exposure time of 5 ms. A detailed description of the setup can be found in previous work (*20*, *39*, *74*) Briefly, for the measurements reported in this study, a 3 W, 532 nm DPSS laser (Del Mar Photonics) is used to illuminate the device through a water-immersion objective (UPlanSApo 60x W, 1.2 NA, Olympus). The emission is filtered with a 532 nm cut-off dichroic and long-pass filters (Chroma), and imaged on an sCMOS camera (Kinetix, Photometrix).

S2.2. Data analysis to determine the escape time of a molecular species

Recorded movies of molecules in the trap landscape are analysed as described in previous work (*20*). Durations of residence times, Δt, of molecules in traps are extracted from intensity vs. time traces, recorded for each trap locus. At present, a one-minute movie typically yields ca.10^4 escape events. The average escape time is the mean of the all the residence times collected, $t_{\mathrm{av}} = \frac{1}{N}\sum \Delta t$, for a total number of residence times, *N*. The uncertainty on determining t_{av} in such a measurement, $t_{\mathrm{av,e}} \cong \frac{t_{\mathrm{av}}}{\sqrt{N}}$, is calculated as the standard error of all recorded residence times. A histogram of residence times is fitted with a decaying single exponential function ($k = 1$ in Eq. (S1)), e.g., to determine a fitted escape time, t_{esc}. The fit error on t_{esc} captures the standard error of the mean, and in general corresponds well to $t_{\mathrm{av,e}}$ as expected.

Fitting histograms of escape-time data

Molecular residence times, Δt, are identified from the intensity vs. time traces in each trap as described in previous work (*20*). The events are binned into a histogram with n bins, where each bin is centered at $\Delta t_i = t_{\mathrm{exp}} + (i-1)\left(t_{\mathrm{exp}} + t_{\mathrm{lag}}\right)$ where i represents the length of the event in number of frames, and t_{exp} and t_{lag} are the exposure time and lag time from the imaging. Then, event counts per bin (N_i) are normalized relative to the total number of events, N, as $\frac{N_i}{N\left(t_{\mathrm{exp}}+t_{\mathrm{lag}}\right)}$, to give probability density distributions $P(\Delta t_i)$. The normalized error on each bin is defined as $\sigma_i = \frac{\sqrt{N_i}}{N\left(t_{\mathrm{exp}}+t_{\mathrm{lag}}\right)}$, while errors on bins containing zero events are set equal to $\left[N\left(t_{\mathrm{exp}} + t_{\mathrm{lag}}\right)\right]^{-1}$. Weighted non-linear regression is used for fitting such histograms with fit function $P_{\mathrm{f}}(\Delta t)$ given by Eq. (S1). Here k is the number of components, E_i is the amplitude of the component with escape timescale, t_i and c is a constant value capturing an experimental offset,

$$P_{\mathrm{f}}(\Delta t) = \sum_{i=1}^{k} \left(\frac{E_i}{t_i} \mathrm{e}^{-\frac{\Delta t}{t_i}} \right) + c \tag{S1}$$

The first bin with $\Delta t_i = t_{\mathrm{exp}}$ is always excluded from fitting to eliminate additional noise from both poor detection of short events arising from transiently trapped molecules, and from random spikes in background noise. Weights in the fitting process are defined as $w_i = 1/\sigma_i^2$ and errors on the fitted escape times are treated as the standard error of the mean.

Determining number of timescale components characterizing an escape-time histogram

Since the escape-time histograms presented in this work have a large number of events per bin (N_i), the Poisson distribution in each bin population may be assumed to approach a normal (Gaussian) distribution. Consequently, the optimal number of exponents needed for fitting a histogram is determined by finding the maximum probability of the expected reduced χ-squared being greater than the observed value, $P_{\mathrm{DoF}}\left(\chi_{\mathrm{r}}^2 > \chi_{\mathrm{r,o}}^2\right)$. For a histogram of n bins with measured $(\Delta t_i, P(\Delta t_i) \pm \sigma_i)$ and fit $\left(\Delta t_i, P_{\mathrm{f}}(\Delta t_i)\right)$ data points, the observed reduced χ-squared is defined as $\chi_{\mathrm{r,o}}^2 = \chi_{\mathrm{o}}^2/DoF$ (*75*), where χ_{o}^2 is the observed χ-squared defined by Eq. (S2), $DoF = n - c_{\mathrm{p}}$ is the degrees of freedom, where c_{p} is the number of free parameters (E_i, t_i, c_i) in the fit function.

$$\chi_{\mathrm{o}}^2 = \sum_{i=2}^{n} \left(\frac{P(\Delta t_i) - P_{\mathrm{f}}(\Delta t_i)}{\sigma_i} \right)^2 \tag{S2}$$

To assess the reduced χ-squared metric, a simulated multi-exponential distribution was generated using as "ground truth" values, the fit results inferred from the experimental histogram of SAM-IV riboswitch (see main text Fig. 4 and Fig. S6A, B). Fig. S6A illustrates the normalized probability density distributions $P_{\mathrm{n}}(\Delta t)$ vs. Δt (green curves) for $N \approx 4.4 \times 10^4$ recorded escape events of SAM-IV riboswitch (green points), also shown in Fig. 4F, after fitting with Eq. (S1) up to k = 4 components. Similarly, Fig. S6B shows the $P_{\mathrm{n}}(\Delta t)$ vs. Δt (grey curves) for an identical number of simulated escape events (black points) based on the fit escape times t_i and event fractions E_i obtained by fitting a 3-exponential decay to the experimental SAM-IV riboswitch (green points and $k = 3$ curve in Fig. S6A). For every fitted curve in Fig. S6A, B the $P_{\mathrm{DoF}}\left(\chi_{\mathrm{r}}^2 > \chi_{\mathrm{r,o}}^2\right)$ was calculated as in Fig. S6C. The probability of a single exponential fit being in agreement with either the simulated (black) or experimental (green) data was less than 0.05%, while for the bi-exponential it increased to ~20% for both. Maximum agreement at $\sim$85% for simulated and $\sim$57% for experimental results was achieved when using a 3-exponential function. Adding a further component led to a decrease in $P_{\mathrm{DoF}}\left(\chi_{\mathrm{r}}^2 > \chi_{\mathrm{r,o}}^2\right)$ for both the simulated and experimentally measured $P(\Delta t_i)$ distributions. As a result, the reduced χ-squared test suggested that a tri-exponential function is the optimal for fitting both the simulated and experimental histogram for the SAM-IV riboswitch. We note that the fitted escape times t_i from the simulated distribution (grey curves) are in good agreement with the ground truth values (3 red dotted lines, Fig. S6D) used for simulations. The fitted escape times are illustrated as normal distributions with the mean given by t_i, and width given by the error on the fitted value. The table lists the ground truth and fitted event fractions E_i with their fit error, for the 3-exponent simulated distribution.

Although the analysis above focused on histograms with three species, larger N values could enable the inference of a larger number of exponential components in the data. For example, Fig. S6E illustrates a histogram of $N \approx 1.5 \times 10^6$ events (black points) constructed based on the three fit escape times and event fractions obtained from the experimental ground truth values from the SAM-IV riboswitch case, including two additional time components. The

six curves in Fig. S6E represent fits to the simulated data of the form of Eq. (S1), where k = 1, …, 6 and $P_{\mathrm{DoF}}\left(\chi_{\mathrm{r}}^2 > \chi_{\mathrm{r,o}}^2\right)$ is determined for each of the 6 fits, as illustrated in Fig. S6F. The probability of single, double and triple exponential fits agreeing with the simulated data was less than 0.05%, while for the 4-exponential it increased to approximately ~18%. Best agreement at ~65% was achieved when using a 5-exponential function. Adding a sixth exponent led to a decrease in $P_{\mathrm{DoF}}\left(\chi_{\mathrm{r}}^2 > \chi_{\mathrm{r,o}}^2\right) \approx 55\%$, with the additional component notably duplicating the t_i value from the fifth exponential. The analysis thus shows that the reduced χ-squared test suggested a 5-exponential function provides an optimal fit to the simulated data capturing the known ground truth. Fig. S6G illustrates the good agreement between the average fitted escape times t_i from the simulated distribution (grey curves) and the ground truth values (5 red dotted lines) used in the simulation. These fitted escape times are illustrated as normal distributions with the mean given by the average t_i from 5 repetitions of the simulation, and width given by the standard deviation of fitted mean values. The table lists the ground truth and average simulated event fractions (E_i) with the quoted error given as the standard deviation of 5 repetitions of the simulation.

S2.3. Nanoslit surface passivation procedure

Prior to measurements of proteins, the devices are often, but not always, passivated to reduce non-specific binding of molecules to surfaces (sticking) following procedures described by Emilsson et al. (*76*). The device is first flushed with Piranha solution (3:1 Sulfuric acid and hydrogen peroxide, Sigma-Aldrich) for 20 min, followed by 20 min of 1 M KOH (Sigma-Aldrich). After a thorough clean with DI water, the device is then incubated with mPEG-silane (5 kDa MW, Sigma-Aldrich) prepared in a solution of 1 mL ethanol, 5 µl 10 mM acetic acid (both Sigma-Aldrich) and 25 µL water in a dark room for at least 48 hr. Then the device is cleaned with water for 5 min, and 0.5% Tween 20 in PBS (pH 7.4) for 20 min. To assess the change in slit height following the coating, the slit height was measured both before and after coating using the calibration procedure relying on four globular proteins, as outlined in section S6.2, data is shown in Fig. S5. For example, for two devices of the same nominal fabricated height ($h_1 = 40$ nm) we found that $h_1 = 47.5 \pm 1.2$ nm without coating and $h_1 = 27.7 \pm 0.4$ nm with mPEG coating as shown in Fig. S5C. This is consistent with a 5 kDa mPEG brush thickness of ca.10 nm as reported by Emilsson et al. (*76*). 0.03% Tween 20 was sometimes added to the measurement buffer to further facilitate the device passivation and minimize non-specific adhesion.

S2.4. High-throughput single-molecule counting down to femtomolar concentrations

In some instances, device passivation was not sufficient to completely reduce the sticking, leading to lower than expected number of escape events detected in our experiments. Note that here we compare counts with those expected for highly charged DNA molecules that are known to display minimal adhesion to the slit surfaces. In general, we expect sticking to surfaces occurring in any part of the experimental protocol to affect the number of recorded counts. E.g., these differences in observed counts could also be due to sticking to preparation tubes, despite the use of LoBind tubes (Eppendorf) for sample preparation. To quantify this reduction in concentration a series of reference measurements of molecules which showed minimal signs of sticking were performed. The nominal concentration of 5 nt single-stranded DNA and insulin were measured before preparation at ca. 1 µM using a Nanodrop. Serial dilutions of each sample from 10 fM to 100 pM were prepared in PBS (pH 7.4) with 0.03% Tween 20 added. The samples were loaded into a freshly coated device with h_1≈70 nm, and the counts were measured under flow for 1-20 minutes, depending on sample concentration. The number of escape events per unit time (event rate) were measured for each sample using an identical

microscope field of view, and a flow rate of ca. 100 μm/s. The flow rate, is estimated from the volume flow rate, Q, using the Hagen-Poiseuille equation for a rectangular channel, Eq. (S3) (*77*), for a pressure drop ΔP, across a channel of length, L, filled with fluid of viscosity, η, and assuming the channel width, W' is much larger than the height.

$$Q \approx \frac{\Delta P\, h_1^3\, W'}{12\eta L}\left(1 - 0.63\,\frac{h_1}{W'}\right) \tag{S3}$$

A linear fit of log(concentration) vs log(event rate) recovers the quantitative sensitivity of ETsy in the current configuration to molecular concentration down to low femtomolar concentrations – a "Femtodrop" (Fig. S5D)! Increased throughput and lowering of this detection threshold by at least a factor 100 may be achieved through further optimization of device design. This approach provides an excellent measure of concentration at the ultra-low concentrations inaccessible to other low-sample volume laboratory measurement techniques such as the Nanodrop and is used to estimate actual measured concentrations of proteins in affinity measurements in section S8. The error bar on each data point in (Fig. S5D) is calculated as (event rate) $/\sqrt{N}$, for a total number of events, N.

S3. Theoretical model relating average escape time to molecular properties r_{H} and D_{s}

The average residence time of molecules in traps, which we refer to as the escape time t_{esc}, may be written as follows:

$$t_{\mathrm{esc}} = t_{\mathrm{r}} \exp\,(W/k_{\mathrm{B}}T) \tag{S4}$$

where W is the free energy difference between molecular states in the pocket and the slit, T is the absolute temperature and k_{B} is Boltzmann's constant. Furthermore, $t_{\mathrm{r}} \propto \frac{{D_{\mathrm{p}}}^2}{16 D_{\mathrm{t}}}$ denotes an estimate of the residence time of the molecule in a pocket region of diameter D_{p}, in the absence of any modulation of slit height (i.e., $h_1 = h_2$). Here D_{t} is the translational diffusion coefficient, which can be related to the Stokes's radius, r_{H}, of a particle in a fluid of viscosity η by the Stokes-Einstein equation:

$$D_{\mathrm{t}} = \frac{k_{\mathrm{B}}T}{6\pi\eta r_{\mathrm{H}}} \tag{S5}$$

where T is the absolute temperature and k_{B} is the Boltzmann constant. In contrast to previous studies using the electrostatic fluidic trap, in this work we operate exclusively at high salt concentrations (~150 mM), which ensures that the electrostatic contribution to W is negligible and the trap is purely entropic (*21*). We therefore have $W \cong \Delta F_{\mathrm{ent}} = -T(\Delta S_{\mathrm{trans}} + \Delta S_{\mathrm{conf}})$. Importantly, the translational entropy contribution for an idealized 'point object', with no spatial extent, may be given by $\Delta S_{\mathrm{trans}} = -k_{\mathrm{B}}\ln\,(h_2/h_1)$, whilst the second, configurational entropy term, ΔS_{conf}, which may enter the picture for large non-spherical objects is discussed later.

In the case of low confinement of the escape process, such as for small hard spheres whose radius $R \approx r_{\mathrm{H}} << h_1$, we may set $\Delta S_{\mathrm{conf}} = 0$ in Eq. (S4) which thus reduces to Eq. (S6) as mentioned in the main text.

$$t_{\mathrm{esc}} \propto r_{\mathrm{H}}(h_2/h_1) \tag{S6}$$

In the case of greater nanoscale confinement such as for larger spheres, or shallower slits, where h_1 approaches the diameter of a sphere, D (given by the regime $D < h_1 \ll h_2$) we expect a non-linear enhancement of escape times with increasing object size D, see Fig. S7. The reason

behind the enhancement lies in the ratio of the effective number of translational states accessible to a finite-sized molecule in the pocket vs. in the slit. This quantity maybe expressed more accurately as a function of object diameter, i.e., by the ratio $\frac{h_2-D}{h_1-D}$, which now supersedes h_2/h_1 for a point object in Eq. (S6), and reveals how the diameter, D, of the object can play an even strong role in the experimental readout. Furthermore, a measure of the radius of the object continues to appear in t_r. Since we generally always have $D \ll h_2$, Eq. (S6) now reads

$$\begin{aligned} t_\mathrm{esc} &\propto r_\mathrm{H}(h_2/h_1)\,(1-D/h_1)^{-1} \\ &= r_\mathrm{H}(h_2/h_1)[1 + D/h_1 + (D/h_1)^2 + (D/h_1)^3 + \cdots \end{aligned} \tag{S7}$$

When D is a substantial fraction of the slit height h_1, a large number of terms in Eq. (S7) contribute to t_esc, which therefore increases strongly with the diameter of the object, offering the prospect of greatly enhancing disparities in measured escape timescales for similar molecules (see Fig. 2C of main text and Fig. S7). For example, in the regime of strong confinement given by, e.g., $D/h_1 = 0.5$, the 4th order term contributes 3% to the measured value of t_esc which is much larger than our typical measured precision of <1%. This showcases the power of h_1 as a tunable experimental parameter that directly impacts the sensitivity of the readout to small changes in molecular diameter. Throughout this work $260 < h_2 < 360$ nm, and note that because $h_2 \gg D$ the measured response is not very sensitive to small changes in h_2, in contrast to h_1.

$$t_\mathrm{esc} \propto r_\mathrm{H}\left(\frac{h_2 - D_\mathrm{s}}{h_1 - D_\mathrm{s}}\right) \tag{S8}$$

The response expected for $D < h_1$, is given by Eq. (S8) which is used to interpret all escape times in this study. For globular molecules with hydrodynamic radius, r_H, we use have bounding sphere diameter, $D_\mathrm{s} = D = 2r_\mathrm{H}$. For non-spherical molecules we use $D = D_s$ in Eq. (S8) as discussed in the main text which accounts for the rotational (orientational) contribution to the entropy difference. In the regime of extreme of extreme confinement given by $D \geq h_1$, further conformational entropy contributions enter the picture but this regime is not considered in the present study.

S4. ETsy measurements of biomolecules

Fluorophore-derivatives, proteins and nucleic acids (DNA) shown in Fig. 2A of the main text were all measured on the same device with $h_1 \approx 25$ nm.

S4.1. Proteins

Insulin (INS), ubiquitin (Ub), thioredoxin (TRX), RNase A, myoglobin (MB), β-lactoglobulin (LBG), carbonic anhydrase (CA), apo-Transferrin (TF), human insulin detector antibody (IgG), apoferritin (FER), Starmaker-like (Stm-l), Prothymosin α (ProTα) and the proline-rich polypeptide (prepared, where applicable, as outlined in section S1.1) are measured at 1-10 nM concentration in PBS (pH 7.4) with 0.03% Tween 20. The escape time histogram was fitted with a single exponential distribution for all the proteins apart from apoferritin and an escape time was determined. Apoferritin data was fitted with a bi-exponential distribution to accommodate a small fraction of ferritin monomer in the sample, as well as the expected 24-mer (*78*), see Fig. S8. The biexponential fit yielded $t_1 = 16.5$ ms, corresponding to a mass of 26 kDa which is consistent with the presence of some apoferritin monomeric subunits with $M \approx 20$ kDa. t_2 is the timescale plotted in Fig. S8A and corresponds to the 24-mer apoferritin

complex. The globular proteins in main text Fig. 2A are fit with a line $t_{\mathrm{esc}} = c_1 MW^{1/3} + c_2$. While the disordered proteins are fit with a line $t_{\mathrm{esc}} = c_1 MW^{3/5} + c_2$, see Fig. S8B.

S4.2. Dye derivatives

Five chemical derivatives of ATTO532 and ATTO542 are measured at 1 nM concentration in 100 mM NaCl, 0.1 mM Tris (pH 7, Carl Roth). Movies at a frame rate of 200 Hz, with t_{exp} = 3 ms, are recorded for each sample, and escape times determined from a single exponential fit to the escape time histogram. The ability of the measurement to distinguish between small differences in molecular weight is illustrated by a linear fit of t_{esc} vs. molecular weight for the four smallest modifications ($M = 650 - 960$ Da). The mass sensitivity of the readout is estimated by the ratio of minimum uncertainty on the fitted escape times to the slope of the fitted line. This analysis suggests a measurement sensitivity of 8.8 kDa or less than a C-atom. An additional dataset for ATTO542 and its derivatives is displayed in Fig. S8C.The mass sensitivity of the measurement is similar for both ATTO542 as ATTO532.

The short measurement times enable tracking of reaction progress. The NHS-ester modified ATTO532 in a solution of 100 mM NaCl, 0.1 mM Tris (pH 7) was measured repeatedly over the course of 45 minutes, at each timepoint a 1-minute movie was recorded and average escape time, t_{av}, determined from a simple average of all the residence time events of duration Δt. No significant change with time was observed. The experiment was repeated in 100 mM NaCl, 2 mM Tris (pH 9), and a substantial time-dependent decrease in t_{av} was observed. Finally, t_{av} reached a plateau value consistent with that of the carboxylic-acid derivative, consistent with the hydrolysis of the NHS-ester. The half-life, $t_{1/2}$, of NHS ester or the inverse rate constant of the hydrolysis reaction was estimated by fitting the data to an exponential decay

$$t_{\mathrm{av}} = t_{\mathrm{av},0} \exp\left(-t/t_{1/2}\right) + \mathrm{const.}$$

Where $t_{\mathrm{av},0}$ denotes the average escape time measured at the start ($t = 0$) of the time-series.

To emphasize the measurement precision, and the implied ability to detect an overall change in average molecular weight of the sample in a time-series, a series of spectra are displayed in Fig. S8C. Each spectrum is normal distribution constructed using the average t_{av} value of all escape events as the mean of the distribution and the corresponding standard error (not the standard deviation) as the width of the Gaussian function. In general, for improved visualization of the data, we present all our escape time measurements as Gaussian functions where the mean value of the distribution is the nominal value, t_{esc} or t_{av}, and the width of the Gaussian, σ, is the standard error of the measurement, obtained either as the error on the fit parameter t_{esc} when using fits to histogrammed data, or as the standard error of the data when using average values t_{av} .

S4.3. DNA

A range of different length double-stranded DNA (dsDNA) from 30 base pairs (bp) to 59 bp (sequences shown in Table S1), are measured at 1 nM concentration in PBS (pH 7.4). A single exponential fit to the escape time histogram is used to determine the escape time. The measurement was repeated in a second device with $h_1 \approx 70$ nm. The ability of the measurement to distinguish small differences in DNA length is estimated as described in S4.2 and is displayed in main text Fig. 2E.

S5. Independent estimates of biomolecular properties from 3D molecular models and FCS

S5.1. Stokes' radius r_H using HYDROPRO

HYDROPRO was used to make a computational estimate of the translational hydrodynamic radius of various samples, r_H. The program generates a primary hydrodynamic particle model based on atomic coordinate files (e.g., PDB formatted file) by replacing amino acids (or nucleotides) with a collection of overlapping spheres (*30*). This is called the shell model in which the molecular surface is replaced by a shell of smaller overlapping spheres. Note that in this model only the surface exposed to the bulk solvent contributes to the frictional properties and thus the hydrodynamic behavior of the molecules. To compute the r_H for each globular protein reported in main text Figs. 2C, D, we used the respective PDB formatted file ID and the relevant molecular weight, as displayed in Table S4. We assumed a specific volume of 0.73 cm^3/g for all species, although replacing it with values of multiple orders of smaller/larger magnitude did not alter the results.

S5.2. Measured Stokes' radius $r_{H,FCS}$ using Dual-focus fluorescence correlation spectroscopy (2f-FCS)

Dual-focus fluorescence correlation spectroscopy (2f-FCS) was used as an independent measure of effective hydrodynamic radii, $r_{H,FCS}$, of the nucleic acids (dsDNA, dsRNA) presented in Table S5. As described previously a homebuilt dual-focus confocal epi-fluorescence microscope was used to record fluorescence signals from labelled molecules freely diffusing in solution (*34*, *36*). Briefly, two identical 520 ± 15 nm laser beams with a 5 ps duration (NKT Photonics EXB-6 High Power Supercontinuum White Light Laser, in conjunction with a NKT Photonics SuperK Varia wavelength selector), pulsed alternatively at 20 MHz, are separated by a polarizing beam-splitter (Thorlabs), with a 25 ns temporal delay imposed on one arm via ~7.5 m of optical fibre. A Nomarski prism (Olympus) is used to alter the path of the two beams depending on their polarization. Following the objective (UPlanSApo 60x W, 1.2 NA, Olympus), two foci are formed which are separated by a small lateral gap ($w_0 = 335$ nm).

Once molecules are excited, emitted light passes through the same objective, prism, dichroic mirror and is focused on a single pinhole. Then, fluorescence is divided by a 50/50 beam splitter and detected by two single photon avalanche diodes (Micro Photon Devices, band-passed using filters from Chroma, Thorlabs and Semrock). Single photons from both detectors are independently recorded in time-tagged time-resolved (TTTR) mode which allows the calculation of both the auto-correlation functions (ACF) for each focus separately and the cross-correlation function (CCF) for photons from the two foci. Cross-correlation yields a timescale for diffusive transport of a molecular species between the optical foci, and hence a characteristic diffusion coefficient, D_t, for the species, as outlined by (*79*). The hydrodynamic radius, r_H is calculated from D_t using the Stokes-Einstein Eq. (S5).

The diffusion coefficient is quantified by statistical bootstrapping which involves averaging over several fitting repetitions of the CCF, each with a different random subset of all recorded data. The temperature T is determined by the precision of the lab thermometer (0.1°C). Each molecule listed in Table S5 was measured 3 times (freshly prepared sample per run) for 20 min. $r_{H,FCS}$ and its error were calculated by taking the average and standard error of the mean, respectively, from these three repeats. Measurement uncertainty in $r_{H,FCS}$ was attributed to both measurement uncertainty in the diffusion coefficient and temperature, which influences $r_{H,FCS}$ both explicitly through the Stokes-Einstein equation and implicitly through its effect on viscosity. All measurements were conducted on 30 µL of sample at 1-10 nM concentration of the labelled species in 100 mM NaCl and 0.2 mM Tris (pH 7.5). Note that

24x24 mm, #1.0 cover glasses (Menzel – Gläser) were used for all experiments and although the per measurement precision in 2f-FCS can be high (~2-5% imprecision), measurement accuracy as reflected in the standard deviation of fitted r_{H} values from repeated measurements can be significantly poorer.

S5.3. Theoretical estimates of r_{H} for the nucleic acid double helix

In order to determine the hydrodynamic radius of dsDNA and dsRNA, we model the double helix as a cylinder of length l, and axial or aspect ratio $p = \frac{l}{w}$, where w is the cylinder width. We then determine the hydrodynamic radius as follows.

For a cylinder, the translational diffusion coefficients along the long, principle k-axis, and symmetric i, j-axes are given by (*80*, *81*):

$$D_{\mathrm{t,i}} = D_{\mathrm{t,j}} = \frac{k_{\mathrm{B}}T}{4\pi\eta L}\left(\ln p + v_{\parallel}\right) \tag{S9}$$

$$D_{\mathrm{t,k}} = \frac{k_{\mathrm{B}}T}{2\pi\eta L}\left(\ln p + v_{\perp}\right) \tag{S10}$$

Where $v_{\parallel}$ and $v_{\perp}$ are end-effect corrections, and were determined empirically for $2 < p < 20$ by Tirado et al. (*81*) and interpolated for ease of use by De La Torre et al. (*82*):

$$v_{\perp} = 0.839 + \frac{0.185}{p} + \frac{0.233}{p^2} \tag{S11}$$

$$v_{\parallel} = -0.207 + \frac{0.980}{p} - \frac{0.133}{p^2} \tag{S12}$$

The spherically averaged macroscopic diffusion coefficient in a random direction is simply the average of the diffusion coefficients in each direction (*83*):

$$D_{\mathrm{t}} = \frac{1}{3}\left(D_{\mathrm{t,i}} + D_{\mathrm{t,j}} + D_{\mathrm{t,k}}\right) \tag{S13}$$

For cylinders, we thus have (*84*):

$$D_{\mathrm{t}} = \frac{k_{\mathrm{B}}T(\ln p + v)}{3\pi\eta L} = \frac{k_{\mathrm{B}}T}{6\pi\eta r_{\mathrm{H}}} \tag{S14}$$

Where the end-effect term, v is given by (*83*)*:*

$$v = \frac{1}{2}\left(v_{\perp} + v_{\parallel}\right) = 0.312 + \frac{0.565}{p} - \frac{0.100}{p^2} \tag{S15}$$

S5.4. Determination of r_{H} and D_{s} relation for the SAM-IV riboswitch

In order to estimate r_{H} and D_{s} for the riboswitch, or indeed for any molecule, from an ETsy measurement performed for single slit height h_1 we need to include some additional constraint in the analysis, e.g., an independent relationship connecting r_{H} and D_{s} if available. We therefore model the RNA as a spheroid of a fixed volume $V = 4\pi l w^2/3 = 40$ nm^3 which provides a relation between r_{H} and D_{s} as outlined below. As before, we define the unique axis

to be the k-axis, of length l, the equatorial axes (i,j-axes) to be of length w, and the axial ratio is given by $p = l/w$. We adopt the Perrin S-factor notation which describes the diffusion coefficients about the different axes for spheroids as follows (*29*, *85*, *86*):

$$D_{\mathrm{t,k}} = k_{\mathrm{B}}T\left(\frac{4\pi}{3p^2V}\right)^{\frac{1}{3}}\frac{(2p^2-1)S(p)-2p^2}{16\pi\eta(p^2-1)} \tag{S16}$$

$$D_{\mathrm{t,i}} = D_{\mathrm{t,j}} = k_{\mathrm{B}}T\left(\frac{4\pi}{3p^2V}\right)^{\frac{1}{3}}\frac{(2p^2-3)S(p)+2p^2}{32\pi\eta(p^2-1)} \tag{S17}$$

With $S(p)$ defined differently for prolate and oblate spheroids:

$$S(p) = \frac{2p}{\sqrt{p^2-1}}\ln\left(p+\sqrt{p^2-1}\right), \text{for prolate spheroids } (p>1) \tag{S18}$$

$$S(p) = \frac{p}{\sqrt{1-p^2}}\tan^{-1}\left(p^{-1}\sqrt{1-p^2}\right), \text{for oblate spheroids } (p<1) \tag{S19}$$

Using Eq. (S13) for the spherically averaged diffusion coefficient, we obtain a relationship between the long-axis $D_{\mathrm{s}} = l$ and r_{H} as follows:

$$D_{\mathrm{t}} = \frac{S(p)k_BT}{\eta}\left(\frac{1}{1296p^2\pi^2V}\right)^{\frac{1}{3}} = \frac{k_{\mathrm{B}}T}{6\pi\eta r_{\mathrm{H}}} \tag{S20}$$

By setting the bounding sphere diameter D_{s} to the longest dimension of a prolate or oblate spheroid, l or w, with fixed volume V, Eqs. (S18)(S20) provide a relation between r_{H} and D_{s}, which we use in section S6.5 to model the riboswitch.

S5.5. Determining D_{s} from molecular 3D structures

A custom Python script was used to calculate D_{s} for predicted structures of biomolecules. The "miniball" module (https://github.com/marmakoide/miniball?tab=readme-ov-file) was used to estimate the minimum diameter of a sphere which could fully enclose all the atoms of the predicted structure, as imported from the relevant PDB file. This package was vetted on the synthetic system of the vertices of a tetrahedron of side length a, and indeed returned a minimum sphere diameter of $\approx 2\sqrt{3}\,a$ as expected from the geometry.

S6. ETsy measurements of biomolecular structure

In order to interpret structural parameters from escape time measurements main text Eq. (1) is used, and repeated here for clarity, Eq. (S21). To verify this relation and to determine the parameters A and t_{o} for our experimental design, we ran Brownian Dynamics (BD) simulations of spherical particles.

$$t_{\mathrm{esc}} = Ar_{\mathrm{H}}\left(\frac{h_2-D_{\mathrm{s}}}{h_1-D_{\mathrm{s}}}\right)+t_{\mathrm{o}} \tag{S21}$$

S6.1. BD Simulations on spherical particles

BD simulations were performed in 3D for spherical particles diffusing in a nanostructured parallel plate landscape. The trajectory of the particle is given by its instantaneous coordinates which were averaged over the exposure time t_{exp} to generate an experimentally relevant trajectory from which escape times were determined as described previously (*20*, *21*, *39*). We used translational diffusion coefficients, D_{t}, for a particle of radius, $R = r_{\mathrm{H}}$ as given by Eq. (S5).

A series of different sized spheres were simulated with 0.5 < R < 10 nm to compare with experimental measurements over the same size range. Each series of spheres were simulated in a series of different trap geometries with 10 < h_1< 70 nm, and $d = h_2 - h_1 = 300$ nm, to compare with the range of different device geometries used in experiments in this study. Trajectories consisting of 10,000 escape events were simulated for each particle in each device geometry, and escape times were determined in the same way as experiments. Using Eq. (S22), and setting $2r_{\mathrm{H}} = D_{\mathrm{s}}$, a global fit of these escape times as a function of r_{H} and h_1 yielded a functional form of A as given by Eq. (S22), as shown in Fig. S7, as well as a value for t_{o}. We note that $t_{\mathrm{o}} \approx t_{\mathrm{exp}}$.

$$A = \alpha\left(\frac{h_1}{h_2}\right)^{\beta} + \gamma, \tag{S22}$$

We obtain fit values of $\alpha = 0.16 \pm 0.34$ s/nm, β =-2.35±0.60 and γ =0.265±0.062 s/μm, which means A is only weakly dependent on h_1. For h_1/h_2 values in this study, A varies between 0.27-0.33 s/μm. As mentioned in section S3, the product, Ar_{H}, can be thought of as the residence time, t_{r}, of a particle in a pocket region in the absence of any modulation of slit height, i.e., $h_1 = h_2$. This residence time can be estimated by the diffusion time of a particle to move a distance s in a 2D plane, $t_{\mathrm{r}} = \frac{s^2}{4D_{\mathrm{t}}}$, where we might expect $s$to equal the pocket radius. For a typical value of $A = 0.3$ s/μm in this work, $s = 443$ nm, which is similar to the pocket radius of 275 nm. Indeed, we expect this value to exceed the radius, in line with the escape radius, R_{esc}, we impose on the simulated trajectories, as in Ref. (*39*). For this study we used an escape radius of the following form, $R_{\mathrm{esc}}(\mathrm{nm}) = 18718(r_{\mathrm{H}} \times 10^9)^{-0.0037} - 18051 \approx 600$ nm. These parameters fully define the governing Eq. (S21) and are used to obtain molecular size information from escape-time measurements in this study.

S6.2. Calibration of the measurement using globular proteins

INS, TF and Apoferritin were measured as outlined in section S4.1. r_{H} for the proteins are calculated using HYDROPRO as outlined in section S5.1. Using Eq. (S21), h_1 was determined using α, β, γ and t_{o} parameters determined from simulations as described in section S6.1, and $h_2 - h_1 = d$ as measured using Atomic Force Microscopy (AFM) scans of the nanostructured surfaces prior to substrate bonding, see main text Fig. 2C. We refer to this procedure as device calibration. Using the calibrated h_1, six further proteins (Ub, TRX, RNase A, MB, CA, LGB) were measured, and their hydrodynamic radii inferred from Eq. (S21), assuming the proteins are globular and therefore that $2r_{\mathrm{H}} = D_{\mathrm{s}}$. We obtained remarkable agreement between measured r_{H} values and the theoretical values for this test set of proteins, see main text Fig. 2D.

For further structural measurements in this study, devices were calibrated similarly using measurements on four globular proteins (INS, CA, TF, FER)

S6.3. Structural measurements on double-stranded nucleic acids

Average escape times for dsDNA and dsRNA of various lengths were measured as outlined in section S4.3 (sequences shown in Table S1). The device was calibrated using the four-point calibration method outlined in S6.2. The increased sensitivity of the escape time to non-spherical objects, compared to globular proteins with the same r_{H}, is illustrated by these measurements as shown in Fig. S10. For prolate objects we expect $D_{\mathrm{s}} \gg 2r_{\mathrm{H}}$ which considerably impacts t_{esc} as shown in Eq. (S21).

We modelled DNA and RNA as rigid cylinders of radii 1.05 nm and 1.18 nm, and of lengths 0.34 nm/bp and 0.25 nm/bp respectively, and determine the r_{H} in each case using the equations given in section S5.3. Note that we also measured the $r_{\mathrm{H,FCS}}$ of our nucleic acids using 2f-FCS, and found good agreement between the experimental and calculated values, see Table S5. Using calculated r_{H} values and assuming $D_{\mathrm{s}} = bn_{\mathrm{bp}}$ we fit Eq. (S21), to determine both A and b values for both DNA and RNA (Fig. S10), where b represents the rise per base pair of the double helix. The fits yielded values for b that are remarkably consistent with the rise per base pair values expected for the B and A-helices characteristic of dsDNA and dsRNA respectively (*34*, *40*). We may attribute the differences in fitted A values from that used for proteins to possible differences in the photophysical properties of the dye when conjugated to a DNA, RNA and amino acid.

For further measurements on DNA in this study (see section S6.4), devices were calibrated using a measurement of different lengths of DNA (15-60 bp, additional sequences shown in Table S1) and fitted using Eq. (S21) with D_{s}=bn_{bp}, to determine h_1 and A for a given device.

S6.4. Structural measurements on DNA nanostructures

The two DNA nanostructures (synthesized as outlined in section S1.2) were measured at 0.1-1 nM concentration in 100 mM NaCl 0.2 mM Tris (pH 7.5) in three devices, each with a different geometry. 0.1 mM $MgCl_2$ was included for measurements of the bundle, but not for the square-tile. A single exponential fit to the escape time histogram is used to determine the average escape time t_{esc}. To measure the size and shape, r_{H} and D_{s}, of the two species, we considered the two measurements in the shallower slits (devices *a* and *b*) as smaller h_1 gives greater sensitivity to D_{s}. Each device was calibrated using a series of different lengths of dsDNA as outlined in section S6.3 and Fig. S11A. For devices with deeper slits, the value of h_1 in the governing Eq. (S21) was fixed to the value measured using AFM rather than from DNA-based calibration. This is because when D_{s}<<h_1 the calibration procedure does not deliver accurate measurements of slit height determining the device geometry. According to Eq. (S21), a single measurement gives a curve of possible (r_{H}, D_{s}) values describing the molecule. By considering two independent measurements of the same molecule in devices with slits of different height h_1 we can simultaneously solve for r_{H} and D_{s}. Since repeated measurements on t_{esc} are normally distributed with a standard deviation given approximately by the fit error we obtain a corresponding map of a probability distribution of possible solutions in each case. An intersection of these distribution for two independent measurements yields a 2D distribution on a D_{s} vs r_{H} plot that may be fitted with a 2D Gaussian to infer r_{H} and D_{s} values, and corresponding uncertainties, see main text Fig. 3C. This approach ignores the potential inaccuracy of the calibration of h_1 for each device, and of A. Accounting for these inaccuracies we obtain the probability map of solutions Fig. S11B. This is generated by selecting a pair of heights, $h_{1,a}$ and $h_{1,b}$, for devices *a* and *b* respectively, from normal distributions following their respective uncertainties from calibration. We then select escape time values for $t_{\mathrm{esc,tile}}$ and $t_{\mathrm{esc,bundle}}$ drawing from normal distributions about the mean values from the measurements that follow their respective measurement uncertainties. The unique solutions to these pairs of values are calculated, $D_{\mathrm{s,tile}}$, $r_{\mathrm{H,tile}}$ and $D_{\mathrm{s,bundle}}$, $r_{\mathrm{H,bundle}}$, along

with the differences between nanostructures. This process is repeated 10,000 times and a histogram of D_s-r_H values for each species is produced, along with a histogram of the differences between nanostructures Fig. S11B.

We note that for a given device, the same h_1 and A is used for each species. The result is that although the range of possible r_H and D_s values for each species is rather large, as shown by the probability map Fig. S11B, given a specific h_1 and A value, a highly precise measurement of *differences between the species* emerges as in main text Fig. 3C. We note that our definition of r_H and D_s requires $D_s > 2r_H$; this physical requirement is applied to all displayed probability maps in keeping with our interpretation of these parameters.

To validate our measured molecular dimensions, oxDNA (*44, 87, 88*) was used to simulate 2000 instances of the nanostructures, and D_s was computed for each frame. A series of exemplary snapshots are included in Fig. S11C. The mean and standard error on the mean of D_s for each nanostructure was calculated for comparison with experimental results. r_H was calculated using HYDROPRO as outlined in section S5.1, for a series of representative frames to estimate a mean and standard deviation. These results are displayed in Fig. S11C.

S6.5. Structural inferences on the SAM-IV riboswitch

The SAM-IV riboswitch (synthesized as outlined in section S1.3) was measured at 1 nM concentration in 5 mM $MgCl_2$ PBS (pH 7.4). The escape time histogram was fitted with a multi-exponential distribution, and yielded three distinct timescales t_1, t_2 and t_3, see main text Fig. 4F. In order to make structural inferences on molecular states implied by these different escape times, the riboswitch was modelled as a spheroid, constraining the r_H and D_s to values consistent with the expected volume of such an RNA molecule. Using a typical RNA density of 1.6 g/cm^3 (*89, 90*), and taking the mass of the SAM-VI riboswitch as 38.51 kDa, a range of 40 nm^3 isovolumetric spheroids were considered to constrain the range of possible states. We set the longest axis to D_s since for a prolate spheroid, $l = D_s$. On the other hand, for an oblate spheroid, we have $w = D_s$. Given a fixed volume and using the relation between r_H and D_s outlined in section S5.4, ellipsoids of a given D_s yield r_H values that are unique to the prolate and oblate cases. Each of these permissible (r_H, D_s) combinations yields an expected escape time, t_{esc}, according to Eq. (S21), which may in turn be compared with measurements

The device was calibrated using the four globular proteins as outlined in S6.2, see Fig. S12A. Using Eq. (S21), the measured escape times were mapped onto possible (r_H, D_s) values. Note that the single molecule measurements yield slightly different values for the two conformational timescales and discussed in section S7.3.

To compare our measurement with 3D molecular structural estimates, the r_H and D_s of the published structure for the riboswitch (pdb:6UES, as inferred from cryoEM by Zhang et al. (*47*) was determined as outlined in sections S5.1 and S5.5.

S6.6. Structural measurements on IR-ECD

Labelled IR-ECD (prepared as outlined in section S1.4), was prepared from frozen aliquots by thawing on ice, centrifuging at 10,000 rpm for 10 min at 4^{o}C, and then running the supernatant through a Zeba micro spin desalting columns with 7 K MWCO (Thermo Fisher Scientific). The device was calibrated using four globular proteins as outlined in S6.2, see Fig. S13A. The sample is measured at a nominal concentration of 10 nM labelled IR-ECD in PBS (pH 7.4) and 0.03% Tween 20, both with and without 1 µM Insulin added, in devices with four different geometries. We note that nominal concentrations differ from true concentrations as described later (section S8.10). Prior to measurement the samples were incubated at room temperature for 1 hr. A single exponential fit to the escape time histogram is used to determine the escape time. To measure the size and shape, r_H and D_s, of the Insulin-bound (Liganded),

and free (Apo) samples, we consider two measurements in different height devices, as for the DNA nanostructures in section S6.4. We consider the intersection of r_{H}-D_{s} solution curves from devices '*a'* and '*c'*, (chosen since these geometries provide the largest difference in gradient between the two curves). Since the measured t_{esc} is normally distributed with the standard deviation given by the standard error of the mean (fit error), we obtain a probability distribution of possible solutions for several different pairs of device heights, $h_{1,\mathrm{a}}$ and $h_{1,\mathrm{c}}$ drawn from distributions whose width is given by the calibration error on the slit height, see main text Fig. 7C.

In order to account for the uncertainties on h_1 values, we repeat this analysis for 10,000 pairs of $h_{1,\mathrm{a}}$ and $h_{1,\mathrm{c}}$ values and $t_{\mathrm{esc,Apo}}$ and $t_{\mathrm{esc,Holo}}$ values. We thus obtain a probability density map of pairs of r_{H} and D_{s} values for each conformation, along with a map of the difference in these two inferred parameters between the two molecular states (Apo and Lig.), see main text Fig. 7D. This analysis is conducted in a manner identical to the DNA nanostructures (outlined in section S6.4). We only consider physically realistic solutions in the domain $D_{\mathrm{s}} > 2r_{\mathrm{H}}$.

We repeat this analysis considering each pair of devices in turn as displayed in Fig. S13B. The most precise results are obtained for devices 'a' and 'c' as expected, on account of the largest different in slopes of Eq. (S21) in these two cases.

To compare with our measured values, the r_{H} and D_{s} was calculated for the published structures of the Liganded structure (PDB:6SOF, (*61*)), partially Ins-bound structure (PDB:7QID, (*61*)), and Apo structure (PDB:4ZXB, (*65*)), as outlined in sections S5.1 and S5.5, see Fig. S13C.

S7. Single-molecule ETsy

S7.1. Single-particle tracking and analysis

Ten-fold lower concentrations were used for ease of single particle tracking in single molecule experiments. Movies were analysed using the same method as previously described to identify escape events as described in S2.2. Single-particle trajectories were constructed from the escape events using single-particle tracking. We then determined an average event length, $t_{\mathrm{av}} \approx t_{\mathrm{esc}}$, of each trajectory consisting of N_{hop}, escape events. Histograms of t_{av} from each single particle trajectory were constructed using a bin width corresponding to the average uncertainty on t_{av}. The histogram entries are colour coded according to the number of escape events per trajectory to emphasize regions of the plot where we have the lowest uncertainty on molecular-level escape times.

t_{av} follows an Erlang distribution, $f(t_{\mathrm{av}})$, with shape parameter, $k = N_{\mathrm{hop}}$, and scale parameter, $\lambda^{-1} = \tau/N_{\mathrm{hop}}$. where τ is the average escape time of the species of interest. Thus we have

$$f(t_{\mathrm{av}}, k, \lambda) = \frac{\lambda^k {t_{\mathrm{av}}}^{k-1} e^{-\lambda t_{\mathrm{av}}}}{(k-1)!} \tag{S23}$$

For N_{hop}>20, the distribution is well approximated by a normal distribution with mean, $\mu = k\lambda^{-1} = \tau$, and standard deviation, $\sigma = \sqrt{k}\lambda^{-1} = \tau/\sqrt{N_{\mathrm{hop}}}$. Thus, although we fit Erlang distributions to all the data, for simplicity we quote μ and σ values characterizing the corresponding normal distributions. In general, the histograms were fitted with a sum of multiple Erlang distributions with shape parameter fixed to the average number of hops, $k = \langle N_{\mathrm{hop}} \rangle$, of all the particle trajectories fitted.

$$f(t_{\mathrm{av}}, k, \lambda_i) = \sum_i \frac{{\lambda_i}^k \, {t_{\mathrm{av}}}^{k-1} e^{-\lambda_i t_{\mathrm{av}}}}{(k-1)!} \quad \text{(S24)}$$

The number of components fitted, i, was determined using a peak finder algorithm. In some cases, the fitting algorithm returned several components with identical λ_i values. Here we report parameters corresponding to fits containing non-redundant time components.

S7.2. Validation of single molecule spectrum approach using dsDNA

Three samples of DNA were measured at 0.01-0.1 nM concentration in 100 mM NaCl, 0.2 mM Tris (pH 7.5): pure 24 bp, pure 60 bp and a 1:1 mixture of these two samples (sequences available in Table S1). For the pure samples a single Erlang distribution was returned by the fitting function, while for the mixture we obtained a sum of two distributions with mean escape time values, μ_1 and μ_2, that were in good agreement with the respective pure samples. We note good agreement between the histogram width and fitted distribution width which shows that the measurement uncertainty is indeed governed by the statistical limit that depends on the number of escape events collected from a single molecule ($\propto 1/\sqrt{N_{\mathrm{hop}}}$). We note that at present the number of hops is primarily dominated by the photobleaching time of the molecule. Note also that a measured histogram wider than the statistically expected standard deviation given by $\sigma = \mu/\sqrt{\langle N_{\mathrm{hop}} \rangle}$could be indicative of molecular processes such as conformational fluctuations which contribute to the broadening of the measurement. Full analysis of the single molecule data is shown in Fig. S14.

The detected escape events for the mixture were also pooled and analysed using the ensemble-fitting approach. A bi-exponential distribution was fitted to the escape time histogram, and m_2, the fraction of 60ds DNA in all the fluorescent molecules was determine d using the method outlined in section S8.1. The fitted escape times and molecular fractions are displayed in a probability density plot which closely captures the measured single molecule distribution, see Fig. 4B of the main text. We note good agreement between the escape times returned by the ensemble fitting, t_1 and t_2, and the single molecule fitting, μ_1 and μ_2. Similarly, the molecular fractions of each species obtained from the ensemble method, are in good agreement with the inferred single molecule component fractions: both approaches correctly recover the 1:1 input ratios of the two species. Thus, we have validated the single molecule method as a quantitative way of dissecting mixtures.

S7.3. Measurements on the SAM-IV riboswitch

The SAM-IV riboswitch sample was measured at 0.01-0.1 nM concentration in 5 mM $MgCl_2$ PBS (pH 7.4). Single-molecule trajectories were constructed, and a histogram of t_{av} plotted and fit as described in section S7.1. As a control three additional samples were measured similarly, carboxyl-ATTO532, 30 dsDNA, 119 ssDNA (preparation details in section S1.3). Full analysis of the single molecule data is shown in Fig. S16

The measurement on the fluorophore sample returned a single time component with mean μ_{dye}. The 30bp DNA returned a three-component fit with a low-amplitude component at $\mu_1 \approx \mu_{\mathrm{dye}}$ in agreement with the 'dye control', and suggesting the presence of a small amount of free dye in the sample. The fit also suggested a second low-amplitude peak (8%) at a timescale μ_2, and a third peak at μ_3 accounting for over 80% of the probability density. The 120 ssDNA also returned three distributions with one low amplitude peak at $\mu_1 \approx \mu_{\mathrm{dye}}$, a second small amplitude peak (2%), and the main peak at μ_3 accounting for over 90% of the probability density. We attribute the second small amplitude peak (only consisting of ~2-3 molecules) of both the 30 ds and 120 ss histogram either to errors in trajectory construction from possible

overlap of trajectories arising from a free dye and NA molecule, or to the inclusion of a single longer timescale "sticking" event in a dye-molecule trajectory.

The width of the t_{av} distribution measured for the riboswitch clearly indicated a heterogeneous distribution of states. Indeed, a fit to the data returned a three-component Erlang distribution with means μ_1, μ_2 and μ_3. One distribution agreed with that of the dye-control, i.e., $\mu_1 \approx \mu_{\mathrm{dye}}$.

The ensemble analysis of the riboswitch escape time data also yielded three timescales t_1, t_2 and t_3 which are consistent with μ_1, μ_2 and μ_3. A probability density histogram of t_{esc} shows that compared to the single molecule histogram the short timescales contribute a larger total fraction in the ensemble analysis. This reflects a challenge that arises in constructing single particle trajectories for faster diffusing species which have larger jumps between consecutive events. The analysis at present therefore has a slight bias towards generally slower diffusing particles characterized by longer escape times. This was not evident in the validation mixture of 24 and 60 dsDNA since the timescales were more similar and there was no free dye in the mixture.

Representative single molecule traces were constructed by time-averaging the hops over either 4 or 16 consecutive hops as previously described (*20*). These traces confirm that the heterogeneity of the riboswitch sample arises from different conformational states that are stable over the observation window of 6 seconds. We did not obtain evidence of interconversion between different conformational states observed for any of the 349 molecules measured, see Fig. S15.

S8. Assessing the extent of molecular binding in a bi-molecular reaction

To track the progress of intermolecular binding we are interested in the fractional abundance of the product, m_2. For a simple two reactant binding system, where A + B → AB, and the observable species are the labelled reactant A and the product AB, we have $m_2 = [AB]/([AB] + [A])$.

S8.1. Determining the fractional abundance of the bound-state, m_2

In an ETsy measurement, we track m_2 by labelling only one of the reactants, A for example. Once the reaction has begun, we expect to have two fluorescent species A and AB. We characterize our two fluorescent species by their respective escape times, $t_{\mathrm{esc,A}} = t_1$, and $t_{\mathrm{esc,AB}} = t_2$, and we expect an ETsy measurement to yield a biexponential distribution of escape event lengths, Δt, as in Eq. (S1). From this equation we note that our measurable in an experiment is the fractional number of events from each exponential distribution characterized by the timescale, t_{i} - the 'event fraction', denoted as E_{i}. We can relate the event fraction to the true molecular fraction by Eq. (S25), which when applied to a two-reactant binding system gives Eq. (S26).

$$m_{\mathrm{i}} = \frac{E_{\mathrm{i}} t_{\mathrm{i}}}{\sum_{\mathrm{j}} E_{\mathrm{j}} t_{\mathrm{j}}} \tag{S25}$$

$$m_2 = \frac{t_2 E_2}{t_2 E_2 + t_1 E_1} \tag{S26}$$

This assumes that in an experiment we observe molecules from each species for the same amount of time T' on average, such that we collect T'/t_i events from each species on average. This assumption relies on each species bleaching at the same rate – a reasonable expectation given that by design the same dye is fluorescing on all visible species. To determine E_1, E_2 and

hence m_2 for our two-reactant binding system proceeding to equilibrium, the timescales t_1 and t_2 were fixed in Eq. (S1). In order to do so, t_1 was measured by species A in absence of B, and t_2 was measured on an equilibrated sample containing A in the presence of a large excess of B, ensuring A is entirely in the bound state AB. We note that the overall average escape time of the mixture, t_{av}, is sufficient to assess the amount of bound-state present.

For a two-reactant binding system, we can relate m_2 to t_{av}, as follows:

$$t_{\mathrm{av}} = \frac{\sum_{\mathrm{i}} E_{\mathrm{i}} t_{\mathrm{i}}}{\sum_{\mathrm{j}} E_{\mathrm{j}}} \tag{S27}$$

$$m_2 = \frac{t_{\mathrm{av}} - t_1}{t_{\mathrm{av}}(1 - t_1/t_2)} \tag{S28}$$

The timescales t_1 and t_2 were assumed to be the minimal and maximal measured t_{av} indicative of 'initial' and 'final' states of a time-dependent reaction, or alternatively, the free and 'fully' bound states in a binding equilibrium measurement. This assumes that at the highest concentration of unlabelled reactant in a binding reaction $t_{\mathrm{av}} = t_2$ and therefore $m_2 = 1$. In instances where not all the reactant is bound at completion, such that $m_2 < 1$, the assumption $t_2 \approx t_{\mathrm{av}}$ no longer holds. However, we note that the relative differences between m_2 values calculated using a nominal $t_2 \approx t_{\mathrm{av}}$ are independent of t_2. As a result m_2 can provide a measure of the extent of binding, which may be used to estimate K_{d} values. Fig. S19 shows verification that K_{d} values may be determined using m_2 values calculated using Eq. (S26) and (S28) interchangeably.

S8.2. Inferring the equilibrium constant, K_{d}

The value of m_2 measured at equilibrium for a two-species binding reaction can be used to infer the binding affinity, K_{d}, of a reaction. By varying the concentration of unlabelled reactant B, $[B]_0$, added to a fixed concentration of labelled A, $[A]_0$, we can relate m_2 at equilibrium to the K_{d} of the reaction as follows (*51*):

$$m_2 = m_2^{\max} \frac{[A]_0}{2} \left[[A]_0 + [B]_0 + K_{\mathrm{d}} \sqrt{([A]_0 + [B]_0 + K_{\mathrm{d}})^2 - 4[A]_0[B]_0} \right] + m_0 \tag{S29}$$

$m_2^{\max}$ denotes the m_2 value at which complex formation has saturated for the highest $[B]_0$ concentrations. In practice this is often observed to be less than 1 and varies due to the exact purity of reactant species (*51*).

S8.3. Measurements of on- and off-rate constants in a binding reaction

The on-rate constant, k_{on}, of a reaction can be measured by tracking m_2 as a function of time after initiation of the reaction. The response of m_2 as a function of time is characterized by an equilibrium rate constant, k_{eq}, as shown in the following Eq. (*1*):

$$m_2 = m_2^{\max}(1 - e^{-k_{\mathrm{eq}} t}) \tag{S30}$$

When $[B]_0$ is in large excess of $[A]_0$, we can express k_{eq} in terms of k_{on} and k_{off}, the reaction off-rate constant, as follows (*51*):

$$k_{\mathrm{eq}} = k_{\mathrm{on}}[B]_0 + k_{\mathrm{off}} \tag{S31}$$

k_{off} can be measured independently by tracking m_2 as a function of time after binding is complete. This can be achieved by addition of a high concentration of a labelled 'chaser'

species C to the fully equilibrated binding reaction. C is designed to bind strongly and essentially irreversibly to free B in solution. The result of this is that once an AB complex spontaneously dissociates the likelihood of A rebinding with another B molecule is essentially zero, resulting in a gradual reduction in the amount of AB in solution. The rate of this decrease in m_2 thus yields an off-rate constant k_{off} which may be given by (*51*):

$$m_2 = m_2^{\max} e^{-k_{\text{off}} t} \tag{S32}$$

The binding affinity can be determined from the measured on and off rate constants according to:

$$K_{\text{d}} = \frac{k_{\text{off}}}{k_{\text{on}}} \tag{S33}$$

S8.4. Estimating the overall uncertainty in measurements of m_2

The general formula for the total error f_{e} on quantity f arising from errors in variables x and y propagation on which f depends may be written as follows (*89*):

$$(f_{\text{e}}^2) = \left(\frac{\delta f}{\delta x}\right)^2 x_{\text{e}}^2 + \left(\frac{\delta f}{\delta y}\right)^2 y_{\text{e}}^2 + 2\left(\frac{\delta f}{\delta x}\frac{\delta f}{\delta y}\right)\text{Cov}(x, y) \tag{S34}$$

where x_{e} and y_{e}denote uncertainties on x and y respectively.

Using Eq. (S26) and (S34), ignoring uncertainties on t_1 and t_2, the uncertainty on m_2, $m_{2,\text{e}}$,can be expressed as follows:

$$\left(m_{2,\text{e}}{}^2\right) = \left[\frac{m_2}{E_2}(1 - m_2)\right]^2 E_{2,\text{e}}{}^2 + \left[\frac{m_2{}^2\tau_1}{\tau_2 E_2}\right]^2 E_{1,\text{e}}{}^2 - \frac{m_2{}^4}{(\tau_2 E_2)^2}\left(\frac{2\tau_1 E_1}{\tau_2 E_2}\text{Cov}(\tau_2 E_2, \tau_1 E_1)\right) \tag{S35}$$

Because E_1 and E_2 are not independent ($E_1 = c - E_2$, with an unknown c), the covariance is non zero. We note that theoretically c =1, but because of spurious events in the first bin of the escape time histogram not fitted as mentioned in section S2.2, $E_1 + E_2 \neq 1$. $\text{Cov}(\tau_2 E_2, \tau_1 E_1) = -\tau_1\tau_2 E_{2,\text{e}}{}^2$ or $\text{Cov}(\tau_2 E_2, \tau_1 E_1) = -\tau_1\tau_2 E_{1,\text{e}}{}^2$. We note that theoretically these should be identical, but in practice we get different uncertainties for E_1 and E_2. We therefore use the larger of the two errors to estimate the overall uncertainty, the final expression given as follows:

$$\left(m_{2,\text{e}}{}^2\right) = \left[\frac{m_2}{E_2}(1 - m_2)\right]^2 E_{2,\text{e}}{}^2 + \left[\frac{m_2{}^2\tau_1}{\tau_2 E_2}\right]^2 E_{1,\text{e}}{}^2 + \frac{m_2{}^4}{(\tau_2 E_2)^2}\left(\frac{2\tau_1{}^2 E_1}{E_2}\max\left\{E_{1,\text{e}}{}^2, E_{2,\text{e}}{}^2\right\}\right) \tag{S36}$$

When using t_{av} to determine m_2, using Eq. (S28) and (S34), and ignoring uncertainties in t_1 and t_2, the uncertainty on m_2 is given as:

$$\left(m_{2,\text{e}}\right) = \frac{t_1}{t_{\text{av}}{}^2(1 - t_1/t_2)} t_{\text{av,e}} \tag{S37}$$

S8.5. DNA hybridisation

DNA hybridization provides a carefully controllable and tunable molecular binding system to quantify the ability of ETsy to measure intermolecular binding. We focused on a sequence of ssDNA 11 nt long reported on by Palau *et al*. (*50*). We obtained a series of fluorescently labelled oligos from 7-11 bases in length that represent progressive truncations of the parent 11 nt sequence. We also included a longer fluorescently labelled 15 nt ssDNA sequence. Two unlabelled 200 nt sequences were designed with complementary 5'ends to the

11 ssDNA and 15 ssDNA respectively. The sequences of the 200 nt oligos were optimized to have as few stable secondary structures as possible using the Mfold web server (*92*). The sequences for all DNA used in this study are presented in Table S1.

Equilibrium measurements of affinity constants

To measure the binding affinity, K_{d}, of each short DNA oligo with the complementary 200 ssDNA, a fixed concentration (0.02-1 nM) of oligo, $[A]_0$, was incubated with varying concentrations of complementary 200 ssDNA, $[B]_0$, in 50 mM NaCl Tris-EDTA buffer (pH 8, Sigma). Where possible, $[A]_0$ was chosen to be well below the K_{D} of interest for each oligo. Incubation times, and range of 200ss concentrations added, varied depending on the oligo length. After equilibrium was reached, the fraction of oligo hybridized, m_2, was determined for each reaction by fitting the escape time histogram with a biexponential distribution as outlined in section S2.2. The K_{d} value was determined by fitting m_2 as a function of $[B]_0$, according to Eq. (S29).

Typically, $m_2^{\max} \neq 1$, which we attribute to always having a non-zero fraction of reactant oligo in an inactive form. For ease of comparison between K_{D} values of different length oligos, a normalized m_2 is used such that the $m_{2,\mathrm{fit}} = 1$ at maximal $[B]_0$.

For the slowest reacting mixtures (i.e., with smallest $[B]_0$) equilibrium was deemed to have been reached by an invariance of measured K_{d} with further incubation time, see Fig. S17. For higher $[B]_0$values we incubated the sample for periods much longer than the theoretically expected equilibration time. For several samples the m_2 was also determined using t_{av}, and compared with those determined from the bi-exponential fitting of the escape time histogram, see Fig. S19. The K_{d} values extracted were the same within error, validating that t_{av} can be used as a quantitative measure of the K_{d}.

We note that the time required for equilibration is much longer than expected from literature estimates of diffusion limited on-rates expected for DNA hybridisation (*51*). We attribute this to the possible influence of non-specific association of biomolecules to the walls of the reaction tubes during incubation. This motivated independent measurements using the reaction kinetics approach as a method for K_{d} determination, described next.

Measurements of reaction kinetics

To measure the on-rate constant, k_{on}, of each short, labelled ssDNA oligo at a fixed concentration $[A]_0 = 0.1 - 1$ nM, in 50 mM NaCl Tris-EDTA buffer (Sigma-Aldrich) was loaded into the device. A small volume of complementary 200 ssDNA, at varying concentrations $[B]_0$, was added into the device and sequential measurements were recorded over a duration of 5-180 mins depending on the $[B]_0$. The fraction of oligo hybridized, m_2, was determined as outlined in section S8.3. Eq. (S30) was used to determine k_{eq} for each $[B]_0$. Data for the k_{eq} measurements of 10 ssDNA are shown in Fig. S18A. Because k_{off} is generally very small, a fit of the measured k_{eq} data to Eq (S31) is rather insensitive to this quantity. k_{off} was therefore determined independent in order to infer k_{on} (Eq (S31)).

To measure the off-rate constant, k_{off}, of each short DNA oligo, after the m_2 had plateaued in the k_{eq} measurements, a small volume of high concentration (0.5-5 μM) unlabelled 11 ss, the chaser oligo C, was added to the device and sequential measurements were recorded until m_2 had decayed to close to 0. Eq. (S32) was used to determine k_{off} for each $[B]_0$. Data for the k_{off} measurements of 9, 10 and 11 nt ssDNA are shown in Fig. S18A. An average value was used to determine k_{on} in conjunction with the k_{eq} measurements using Eq. (S31), inset of Fig. S18A.

To explore whether the chaser was facilitating the unbinding of AB and thus influencing the k_{off} measurement, k_{off} was measured for a series of different reaction mixtures, all incubated with constant $[A]_0$ and $[B]_0$ but with varying chaser concentrations $[C]$. k_{off} was determined to be the same within error for the whole range of ratios of $[C]/[B]_0$ measured in this study, thus indicating minimal cooperative unbinding in our system over that concentration range. For higher ratios of $[C]/[B]_0$, a factor two increase in k_{off} was indeed observed, indicative of facilitated unbinding, but this regime is not relevant to our study, see Fig. S18B.

We note that in occasional cases there was a short time delay after loading of the B before any binding was observed. Then after this the expected inverse exponential increase in bound fraction occurred. In these instances, this offset was subtracted from the measurement times to give a more accurate measure of the reaction on rate. Since the response curves are exponential, analysis of any part of the response should yield the same measured rate. We attribute this offset to inconsistent loading times through the microfluidic channels. A similar effect was occasionally observed after addition of the chaser and the offset similarly accounted for.

Calculation of theoretically expected binding affinities

Theoretical K_{d} values were computed for each oligo using the nearest-neighbor Watson Crick parameters as outlined by SantaLucia et al. (*52*).

S8.6. 24bp + EcoRI binding affinity measurement

Labelled 24 bp DNA was measured at 250 pM in PBS (pH 7.4) with varying concentrations of EcoRI added (details of samples in section S1.1). Prior to measurement, the sample was incubated at room temperature for 2 hours. The fraction of oligo bound to the EcoRI, m_2, was determined from fitting the escape time histogram with a biexponential function as outlined in section S8.1. K_{d} for the system was determined using Eq. (S29), with oligo concentration, $[A]_0$, and EcoRI concentration, $[B]_0$. Bar plots in main text Fig. 6A illustrate the relative fractions of unbound and bound oligo for varying concentrations of EcoRI. Repetitions of the affinity measurement are shown in Fig. S20. The m_2 was also determined using t_{av} as outlined in section S8.1, and compared with those determined from the bi-exponential fitting of the escape time histogram, see Fig. S19. The K_{d} values determined were the same within error, further confirming that t_{av} can be used as a quantitative measure of K_{d}.

S8.7. HLA + IgG binding affinity measurement

Labelled HLA (prepared as outlined in section S1.1) was measured at 1 nM in PBS (pH 7.4) with varying concentrations of added unlabelled IgG. Prior to the measurement the samples were incubated at room temperature for 1 hour. The fraction of HLA bound to the IgG, m_2, was determined from t_{av} as outlined in section S8.1. t_{av} was used to track the binding since the HLA is a complex known to dissociate giving a complicated mixture of timescales especially at low nanomolar concentrations (*93, 94*). t_1 and t_2 were determined by the t_{av} measured for the sample with the lowest and highest concentrations of added IgG, respectively. A considerably lower density of molecules was observed in the device than the nominal concentration of sample, measured using the Nanodrop on sample at concentrations that are several orders of magnitude higher. A corrected concentration of 0.1 nM HLA was measured with reference to a 60 bp dsDNA sample as outlined in section S2.4, and was used as $[A]_0$ in Eq. (S29) with the corresponding IgG concentration, $[B]_0$, in order to estimate K_{d} for this system. Repetitions of the affinity measurement are shown in Fig. S20.

S8.8. Insulin aptamer + Insulin binding affinity measurement

A labelled insulin aptamer sequence as reported by Yoshida et al. (*95*) and listed in Table S1, was prepared via annealing in 10 mM $MgCl_2$ (Sigma-Aldrich) PBS (pH 7.4) for 10 min at 95°C, and gradual cooling down at 2°C/min. The folded aptamer was measured at 5 nM concentration in the same buffer with varying concentrations of added insulin. The fraction of aptamer bound to insulin, m_2, was determined from t_{av} as outlined in section S8.1. t_{av} was used to track the binding since the timescales t_1, and t_2 for the unbound and bound aptamer respectively, were within 10% of one another, and therefore the fitting of a biexponential distribution to the escape time histogram was unreliable. K_{d} for the system was determined using Eq. (S29). Repetitions of the affinity measurement are shown in Fig. S21B.

We also note the difference in timescale of the aptamer measured in PBS (pH 7.4) compared to that measured after annealing in 10 mM $MgCl_2$ PBS (pH 7.4). This is consistent with the formation of a G-quadruplex structure, as reported previously (*95*, *96*), which is expected to cause an increase in compaction of the DNA and thus a reduction in escape time. This was also observed in FCS measurements of the unfolded, folded and insulin-bound states, shown in Fig. S21A.

S8.9. Insulin + IR-ECD binding affinity measurement

Labelled insulin (prepared as outlined in section S1.1) was measured at 2 nM concentration in HEPES Buffered Saline (pH 7.2, Merck) with varying concentrations of unlabelled IR-ECD (prepared as outlined in section S1.4). Prior to measurements the samples were incubated overnight at 4°C. The fraction of insulin bound to the IR-ECD, m_2, was determined from fitting the escape time histogram with a bi-exponential distribution as outlined in section S8.1. K_{d} for the system was determined using Eq. (S29) with insulin concentration, $[A]_0$, and IR-ECD concentration, $[B]_0$. We assume that for this measurement of K_{d} the Insulin-IR-ECD is approximately a monovalent binding system, since, despite the four possible insulin binding sites on the IR-ECD, at the rather large concentration ratio of 15:1 [IR-ECD]:[Insulin] corresponding to the mean K_{d} value measured, the likelihood of multivalent Insulin-IR-ECD binding is very low. Repetitions of the affinity measurement are shown in Fig. S20.

S8.10. IR-ECD + Insulin binding affinity measurement

Labelled IR-ECD (prepared as outlined in S1.4 and S6.6) was measured at a calibrated concentration of 0.08 nM (calibrated using 60 dsDNA as a reference molecule as outlined in section S2.4) with varying concentrations of unlabelled insulin. This corresponded to a nominal concentration of 10 nM as measured using Nanodrop before dilution. Prior to measurement, the sample was incubated at room temperature for 1 hr. The fraction of IR-ECD bound to insulin, m_2, was determined from t_{av} as outlined in section S8.1. t_{av} was used to track the binding since the timescales t_1, and t_2 for the unbound and bound aptamer respectively, were within 10% of one another, and therefore the fitting of a biexponential distribution to the escape time histogram was unreliable. The K_{d} for the interaction based on conformation change was calculated using Eq. (S29), see main text Fig. 7E. It has been reported by Gutmann et al. (*61*) that at least 3 insulins were required to bind to the IR-ECD before a significant conformation change occurs. The impact and relative contributions of the underlying processes on the K_{D} estimated from a conformation change is not clear.

This measurement was repeated in simulated human serum (Biochemazone) with 0.03% Tween 20 in solution. We determined K_{d} similarly to in PBS using a calibrated IR-ECD concentration of $[A]_0 = 0.1$ nM, see main text Fig. 7E.

The measurement was also repeated with longer incubation at 4°C overnight. We obtained larger K_d values in this case, which we attribute to non-specific binding of insulin to surfaces of the preparation tubes used for incubation causing a reduction in effective $[B]_0$.

As a control, IR-ECD was measured under identical conditions but with varying concentrations of Epidermal Growth Factor (EGF) (sample preparation as in section S1.1) added instead of insulin, see Fig. S22. Prior to measurement the sample was incubated at room temperature for 1 hr. No change in escape time was observed upon addition of EGF, and hence no conformation change of the IR-ECD was observed, indicating the specificity of the IR-ECD to insulin.

S8.11. Sources of inaccuracy in molecular binding measurements

Accurate measurements of binding affinities are inherently challenging due to inaccuracies in measurement of pico-nanomolar concentrations of the reactants. In general, absorption measurements at low micromolar concentration were used before dilution for measurement. We found considerable variability in these absorption measurements and our measurements therefore must contend with a potential factor 2 uncertainty in quoted concentrations for a given sample. Where a range of concentrations were measured, for example K_d curves with varying $[B]_0$, sequential dilution was used to ensure accurate relative measurement across a K_d curve – here the factor 2 uncertainty in the original stock manifests as a shift in the *x*-axis for all data points. Where a single concentration is measured, as for $[A]_0$, we note that the K_d is relatively insensitive to $[A]_0$ concentration as long as it is well below the K_d. We therefore aimed for this regime for all measurements.

A perhaps more concerning source of concentration inaccuracy is nonspecific binding (sticking) of the sample to preparation tubes and also within the device. Non-stick Eppendorf tubes and device passivation were used to mitigate these issues but in multiple cases we observed considerably lower concentrations of protein than nominally measured before dilution. For fluorescent species we use quantitative molecule counting to calibrate the concentration of freely diffusing molecules.

A final source of inaccuracy in concentration measurements is in determining the concentration of the "active" binding fraction in a sample. We consistently saw reactions complete with <100% binding. We attribute this to a non-negligible 'inactive' fraction of the sample that is unable to bind. This has been widely reported in the literature (*51*), and means that the active $[A]_0$ concentration, $[A]_{0,\mathrm{active}} < [A]_0$. This fraction $[A]_{0,\mathrm{active}}/[A]_0$ can be estimated from the saturation m_2 value at completion of the reaction and is often ~30%. Since the K_D is by design insensitive to $[A]_0$, the $[A]_{0,\mathrm{active}}/[A]_0$ fraction is unimportant, however if generally true, then we may expect a similar fraction of $[B]$ that is similarly inactive, thus suggestive of ~30% inaccuracy in $[B]_0$ and hence in the inferred K_d value.

S9. Supplementary Tables

Table S1 Nucleic acid sequences with colour-coded complementary ssDNA sequences

Sample	Sequence (5' to 3')	Vendor	HPLC purified?
Insulin aptamer	GGT GGT GGG GGG GGT TGG TAG GGT GTC TTC T /3ATTO532N/	IDT	Single
15bp DNA	/5ATTO532N/TAG GAT CCA TAG TAG CTA CTA TGG ATC CTA	IDT	Single
24bp DNA	/5ATTO532N/ TAG GAT CCA TAG AAT TCA GCG TAG /5ATTO532N/ CTA CGC TGA ATT CTA TGG ATC CTA	IDT	Double
30bp DNA	/5ATTO532N/ GGA TGG GAC GGA CCC GGA CAC AGA CAG TGC /5ATTO532N/ GCA CTG TCT GTG TCC GGG TCC GTC CCA TCC	IDT	Double
40bp DNA	/5ATTO532N/ CGC TCA AGG TGG ATG GGA CGG ATT CGG ACA TAG ACA GTG C /5ATTO532N/ GCA CTG TCT ATG TCC GAA TCC GTC CCA TCC ACC TTG AGC G	IDT	Double
50bp DNA	/5ATTO532N/ CGG TCC AGT TGC ACT GTC TAT GTC CGA ATC CGT CCC ATC CAC CTT GAG CG /5ATTO532N/ CGC TCA AGG TGG ATG GGA CGG ATT CGG ACA TAG ACA GTG CAA CTG GAC CG	IDT	Double
56bp DNA	/5ATTO532N/ TA GAA CTA GTG GAT CCC CCG GGC TGC AGG AAT TCG ATA TCG CAA TGT ATA GCA ACT /5ATTO532N/ A GTT GCT ATA CAT TGC GAT ATC GAA TTC CTG CAG CCC GGG GGA TCC ACT AGT TCT A	Microsynth	Double
57bp DNA	/5ATTO532N/ TA GAA CTA GTG GAT CCC CCG GGC TGC AGG AAT TCG ATA TCG CAA TGT ATA GCA ACT T /5ATTO532N/ AA GTT GCT ATA CAT TGC GAT ATC GAA TTC CTG CAG CCC GGG GGA TCC ACT AGT TCT A	IDT	Double
58bp DNA	/5ATTO532N/TAG AAC TAG TGG ATC CCC CGG GCT GCA GGA ATT CGA TAT CGC AAT GTA TAG CAA CTT G /5ATTO532N/ CAA GTT GCT ATA CAT TGC GAT ATC GAA TTC CTG CAG CCC GGG GGA TCC ACT AGT TCT A	IDT	Double
59bp DNA	/5ATTO532N/ TAG AAC TAG TGG ATC CCC CGG GCT GCA GGA ATT CGA TAT CGC AAT GTA TAG CAA CTT GG /5ATTO532N/ CCA AGT TGC TAT ACA TTG CGA TAT CGA ATT CCT GCA GCC CGG GGG ATC CAC TAG TTC TA	IDT	Double
60bp DNA	/5ATTO532N/ TAG AAC TAG TGG ATC CCC CGG GCT GCA GGA ATT CGA TAT CGC AAT GTA TAG CAA CTT GGG /5ATTO532N/ CCC AAG TTG CTA TAC ATT GCG ATA TCG AAT TCC TGC AGC CCG GGG GAT CCA CTA GTT CTA	IDT	Double
30bp RNA	/5ATTO532N/ GGA UGG GAC GGA CCC GGA CAC AGA CAG UGC /5ATTO532N/ GCA CUG UCU GUG UCC GGG UCC GUC CCA UCC	IDT	Double
40bp RNA	/5ATTO532N/ CGC UCA AGG UGG AUG GGA CGG AUU CGG ACA UAG ACA GUG C	IDT	Double

	/5ATTO532N/ GCA CUG UCU AUG UCC GAA UCC GUC CCA UCC ACC UUG AGC G		
50bp RNA	/5ATTO532N/ CGC UCA AGG UGG AUG GGA CGG AUU CGG ACA UAG ACA GUG CAA CUG GAC CG /5ATTO532N/ CGG UCC AGU UGC ACU GUC UAU GUC CGA AUC CGU CCC AUC CAC CUU GAG CG	IDT	Double
60bp RNA	/5ATTO532N/ UAG AAC UAG UGG AUC CCC CGG GCU GCA GGA AUU CGA UAU CGC AAU GUA UAG CAA CUU GGC /5ATTO532N/ GCC AAG UUG CUA UAC AUU GCG AUA UCG AAU UCC UGC AGC CCG GGG GAU CCA CUA GUU CUA	IDT	Double
5ss DNA	/5ATTO532N/TAGAA	IBA	Double
7ss DNA	**TTGTGAG** /3ATTO532N/	IDT	Double
8ss DNA	**GTTGTGAG** /3ATTO532N/	IDT	Double
9ss DNA	**TGTTGTGAG** /3ATTO532N/	IDT	Double
10ss DNA	**C TGT TGT GAG** /3ATTO532N/	IDT	Double
11ss DNA	**CC TGT TGT GAG** /3ATTO532N/	IDT	Double
11ss DNA chaser	**CC TGT TGT GAG**	IDT	No
200ss DNA comp to 7-11	**CTC ACA ACA GG** CTACTTACA ATGAAGATCA GGAACCCCCG GGGTGCAGGA ATTCGATATC GCAATGTATA GCAACTTGGC TAGAAGTAGT GGATCCCCCG GGCTGCAGGA ATTCGATATC GCAATGTATA GCAACTTGGC TAGAAGTAGT GGATCCCCCG GGCTGCAGGA ATTCGATATC GCAATGTATA GCAACTTGGC	IDT	No
15ss DNA	/5Alexa532/ **TAG GTT CCA TAG TAG**	IDT	Single
200ss DNA comp to 15ss	**CTA CTA TGG AAC CTA** TTACA ATGAACATCT CGAACCCCCG GGCTGCAGGA ATTCGATATC GCAATGTATA GCAACTTGGC TAGAAGTAGT GGATCCCCCG GGCTGCAGGA ATTCGATATC GCAATGTATA GCAACTTGGC TAGAAGTAGT GGATCCCCCG GGCTGCAGGA ATTCGATATC	IDT	No
119ss RNA	/5ATTO532N/ GGUCAUGAGU GCCAGCGUCA AGCCCCGGCU UGCUGGCCGG CAACCCUCCA ACCGCGGUGG GGUGCCCCGG GUGAUGACCA GGUUGAGUAG CCGUGACGGC UACGCGGCAA GCGCGGGUC	N/A	Single
119ss DNA	/5ATTO532N/ TATAACTAAT GTATCATCCT AACTGCAGTA ACTCTTTATC GCAATGTATA TCAACTTGTC TAGAACTACC TAATCTCTCG TATTGAATGA ATTTTATATC TTAATGTATA GCAACTTAC	N/A	Single

Table S2 Disordered Protein sequences

Protein	**Sequence**	**Molecular Weight** (kDa)
13-mer proline rich polypeptide	PPPPPPPD PPPPPPPD PPPPPPPD PPPPPPPD PPPPPPPD PPPPPPPD PPPPPPPD PPPPPPPD PPPPPPPD PPPPPPPD PPPPPPP	9.44
ProTα	AHHHHHHSAA LEVLFQGPSD AAVDTSSEIT TKDLKEKKEV VEEAENGRDA PANGNAENEE NGEQEADNEV DEECEEGGEE EEEEEEGDGE EEDGDEDEEA ESATGKRAAE DDEDDDVDTK KQKTDEDC	14.02
Stm-l	GSMKNSDDESKE TSKCAAEVKT TEKTAENPDA TDEPSEETSE GEELDSADDS QDSTDEPSQE TAETKETESC DKSEDKSSAA EETDEEASDS ADVSEAVSES EETSSSPEVR GSKTNNTEDK ISFEEVEDQS DEMSMDKDKK SEESSMRPTD SVKGAEDSAS LESDEMFEES DEQSEEEKST SAPTSEDAGA DSEDESDQSS QDLSKEDDSK GQDAVQSQDD GTTSKADLMD DDLGAEDDKD EEKEEAADGD AKDLGVMIEK DEEEKVEVNK EQVDSNEDQP AKENSDEKDE ADKEKDASDK KDLSMSEEEP DWEEEMGMDD SLEEEEKANK KTSDEKGSFD ETTSDQLDQP DNSTPPADAA M	39.58

Table S3 Details of protein labelling

Protein	**Molar excess** (dye)	**Degree of labelling** (DOL)	**Post-labelling concentration after purification** [µM]
INS	1.5x	0.44	15.5
Ub	2.5x	0.48	177.5
TRX	2.5x	0.36	13.56
RNase A	2.5x	0.55	162.67
MB	2.5x	0.49	115.59
LGB	2.5x	1.08	76.99
CA	2.5x	1.1	24.2
TF	2.5x	2.4	14.5
FER	5x	0.7	22
IgG	5x	3	4.3
HLA	3x	2.2	15.8

Table S4 Measured and structural estimates of Stokes' radii, $r_{\mathrm{H,ETsy}}, r_{\mathrm{H,FCS}}$ and r_{H} of proteins. Molecular weight values include one ATTO532 dye (0.64 kDa), while PDB files used to estimate r_{H} did not take the dye into consideration.

Molecular Weight (kDa) (+ ATTO532)	**Sample**	**PDB ID**	$\boldsymbol{r_{\mathrm{H}}}$ (nm)	$\boldsymbol{r_{\mathrm{H,ETsy}}}$ (nm)	$\boldsymbol{r_{\mathrm{H,FCS}}}$ (nm)
6.44	INS	3I40	1.455	1.32 ± 0.02	1.1 ± 0.2
9.13	Ub	3H1U	1.690	1.46 ± 0.04	1.5 ± 0.2
12.29	TRX	1F6M	1.797	1.83 ± 0.02	1.1 ± 0.1
14.32	RNase A	1FS3	1.954	2.03 ± 0.02	2.2 ± 0.2
17.62	MB	1DWR	2.102	2.17 ± 0.04	1.4 ± 0.1
19.24	LGB	6GE7	2.186	2.29 ± 0.04	2.4 ± 0.1
32.07	CA	4CNX	2.459	2.35 ± 0.06	2.5 ± 0.1
75.77	TF	3QYT	3.647	3.69 ± 0.04	2.8 ± 0.9
465.18	FER	4V1W	6.435	6.68 ± 0.06	4.5 ± 0.1
202.65	IR-ECD (Apo)	4ZXB	5.85	5.70	4.8 ± 0.7
226.08	IR-ECD (Liganded)	6SOF	5.92	5.45	5.1 ± 0.3

Table S5 Measured Stokes' radius, $r_{\mathrm{H,FCS}}$, of fluorescently labelled nucleic acids using dual-focus fluorescence correlation spectroscopy (2f-FCS), and their respective calculated values from cylindrical model (dsRNA & dsDNA) or oxDNA (tile & bundle)

Sample	**base pairs**	$\boldsymbol{r_{\mathrm{H,FCS}}}$ (nm)	$\boldsymbol{r_{\mathrm{H}}}$ (nm)
dsRNA	30	2.13 ± 0.04	2.28
	40	2.38 ± 0.01	2.63
	50	2.63 ± 0.03	2.99
	60	3.04 ± 0.02	3.34
dsDNA	30	2.64 ± 0.05	2.52
	40	3.03 ± 0.05	2.97
	50	3.50 ± 0.03	3.42
	56	3.82 ± 0.04	3.68
	57	4.04 ± 0.02	3.73
	58	3.90 ± 0.04	3.77
	59	4.04 ± 0.05	3.82
	60	3.99 ± 0.09	3.86
Tile	240	6.7 ± 0.2	4.39± 0.08
Bundle	240	6.6 ± 0.1	5.02±0.05

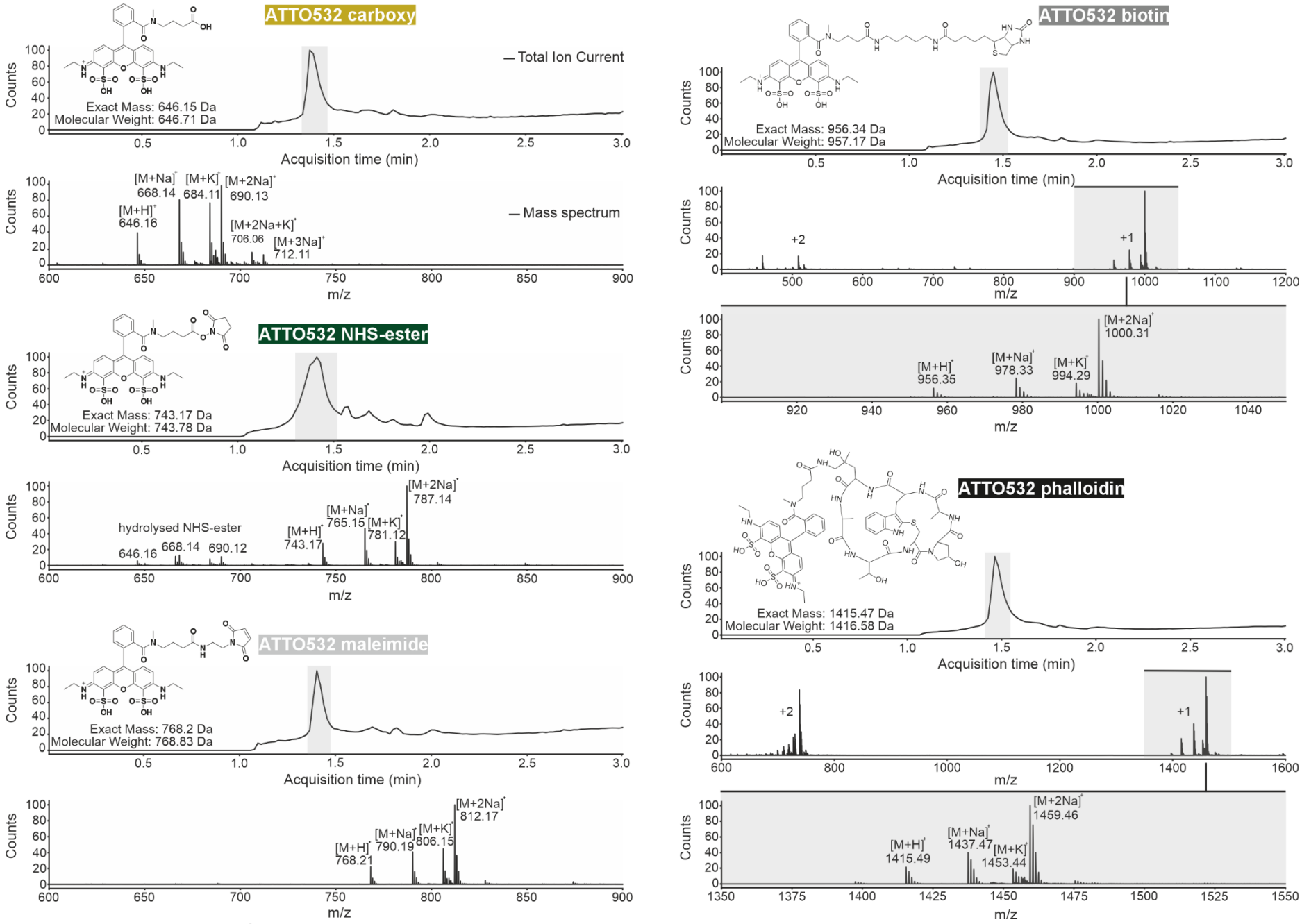


Fig. S1. LC-MS spectra of ATTO532 dye derivatives. Total ion chromatogram (top) and mass spectrum (bottom) are presented in each case.

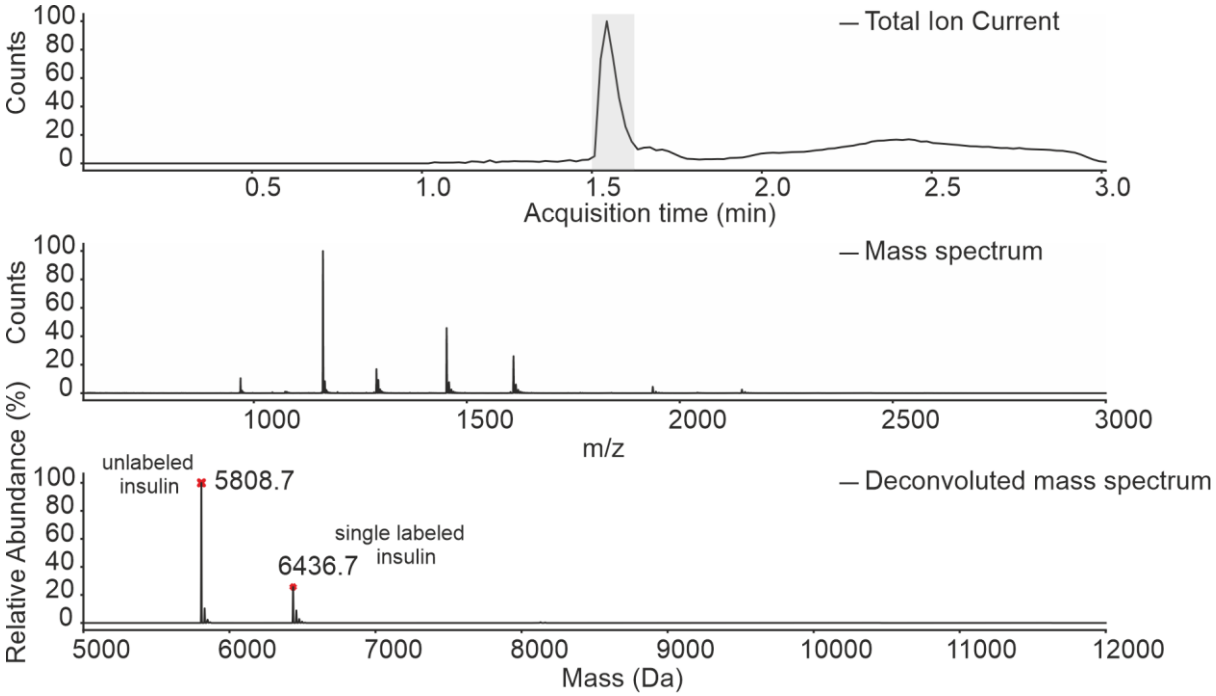


Fig. S2. **Intact mass spectrometry analysis of ATTO532-labelled insulin**. Total ion chromatogram, mass spectrum and deconvoluted mass spectrum are shown. Only unlabelled insulin and insulin carrying a single ATTO532 label were observed.

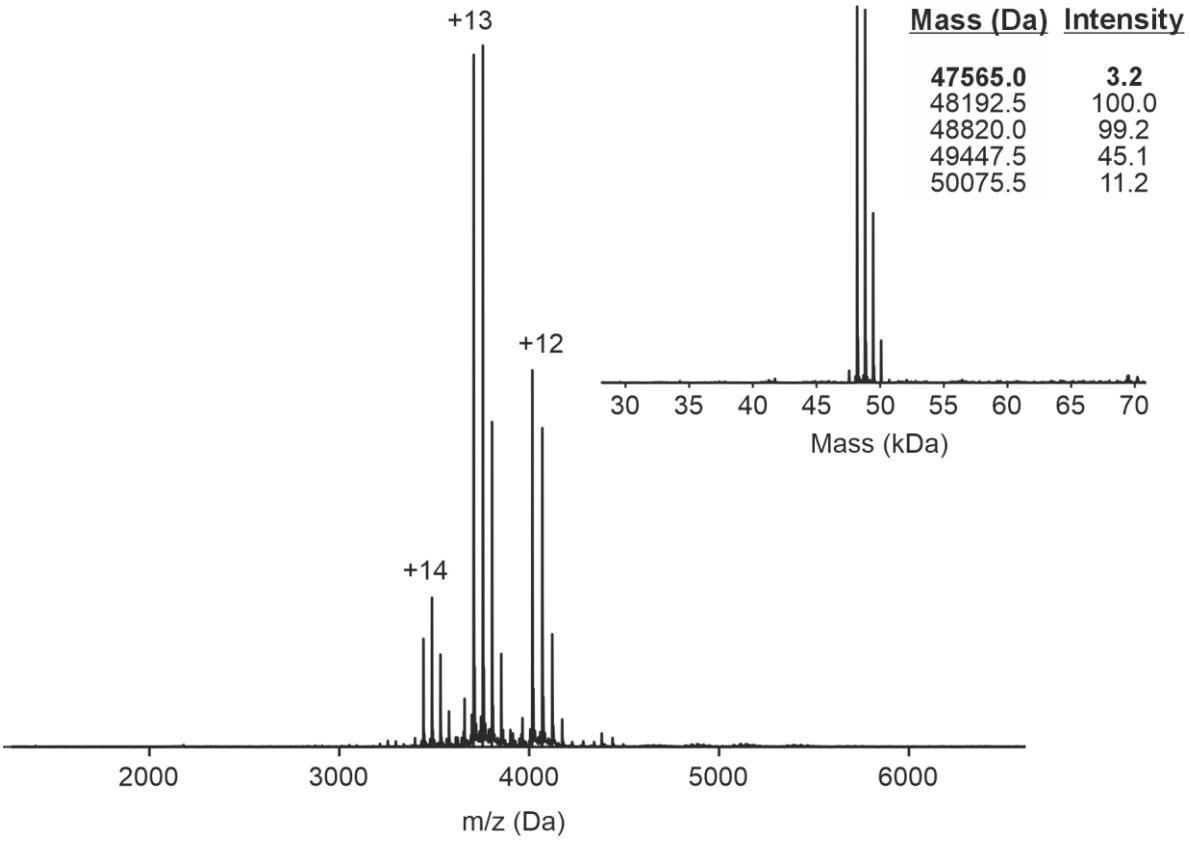


Fig. S3. **Native mass spectrometry analysis of ATTO532-labelled HLA-A*03:01.** The table lists all measured masses. We expect unlabelled HLA to have a nominal mass of 47492 Da which is close to the measured 47565 Da (first table row). Adduct masses of 627.5 Da correspond to a single ATTO532 molecule. The majority of HLA molecules carry one or two labels, but up to four labels were observed on a small fraction (~11%) of the labelled HLA complex.

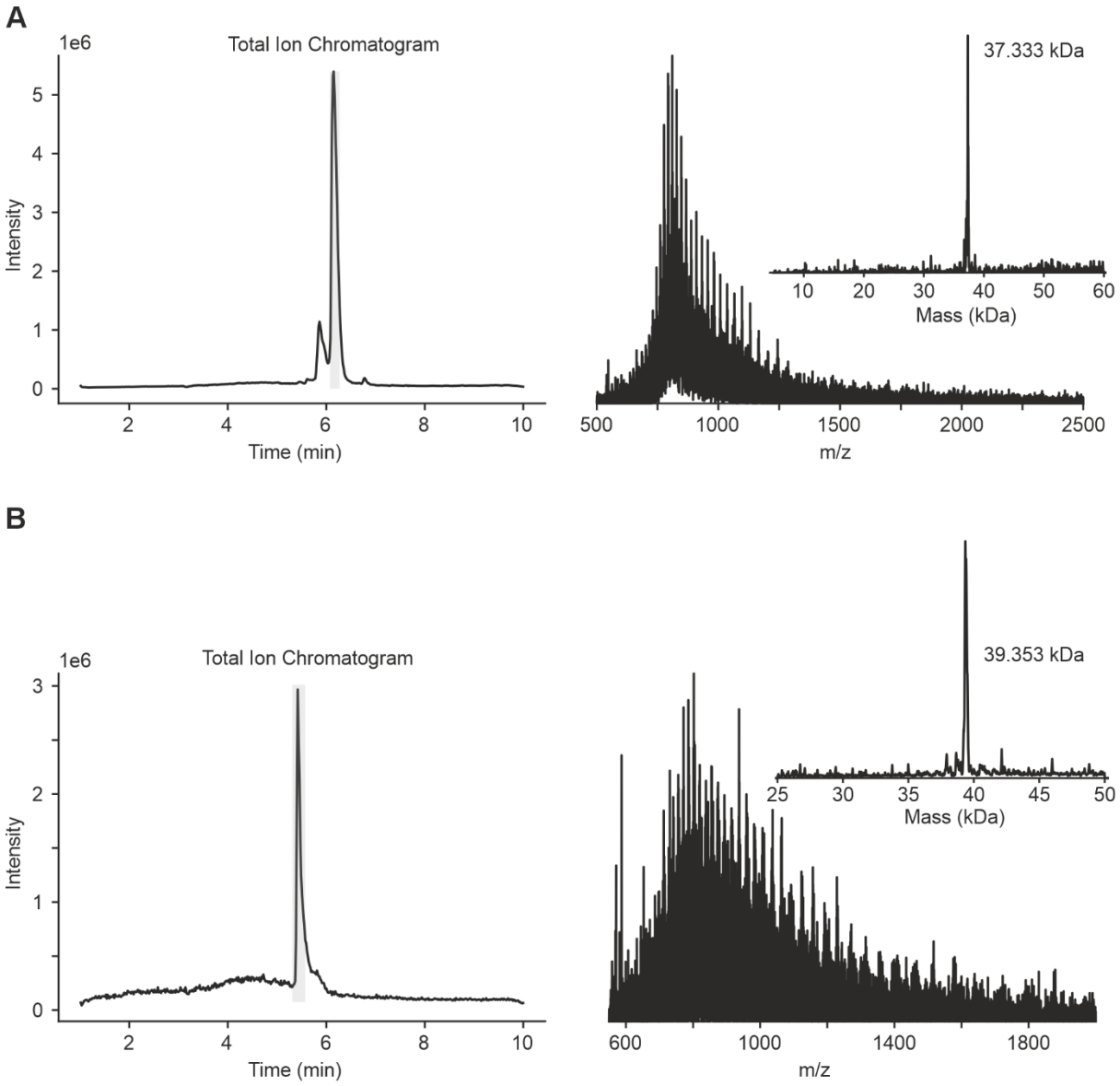


Fig. S4. **LC-MS spectra of (A) 119 nt reference ssDNA and (B) SAM-IV riboswitch.** Total ion chromatogram and mass spectrum are shown in each case. Measured masses agree well with expectations from the sequences: for DNA, $M \approx 37$ kDa, for RNA: $M \approx 39$ kDa. Sequences are shown in S9.Table S1.

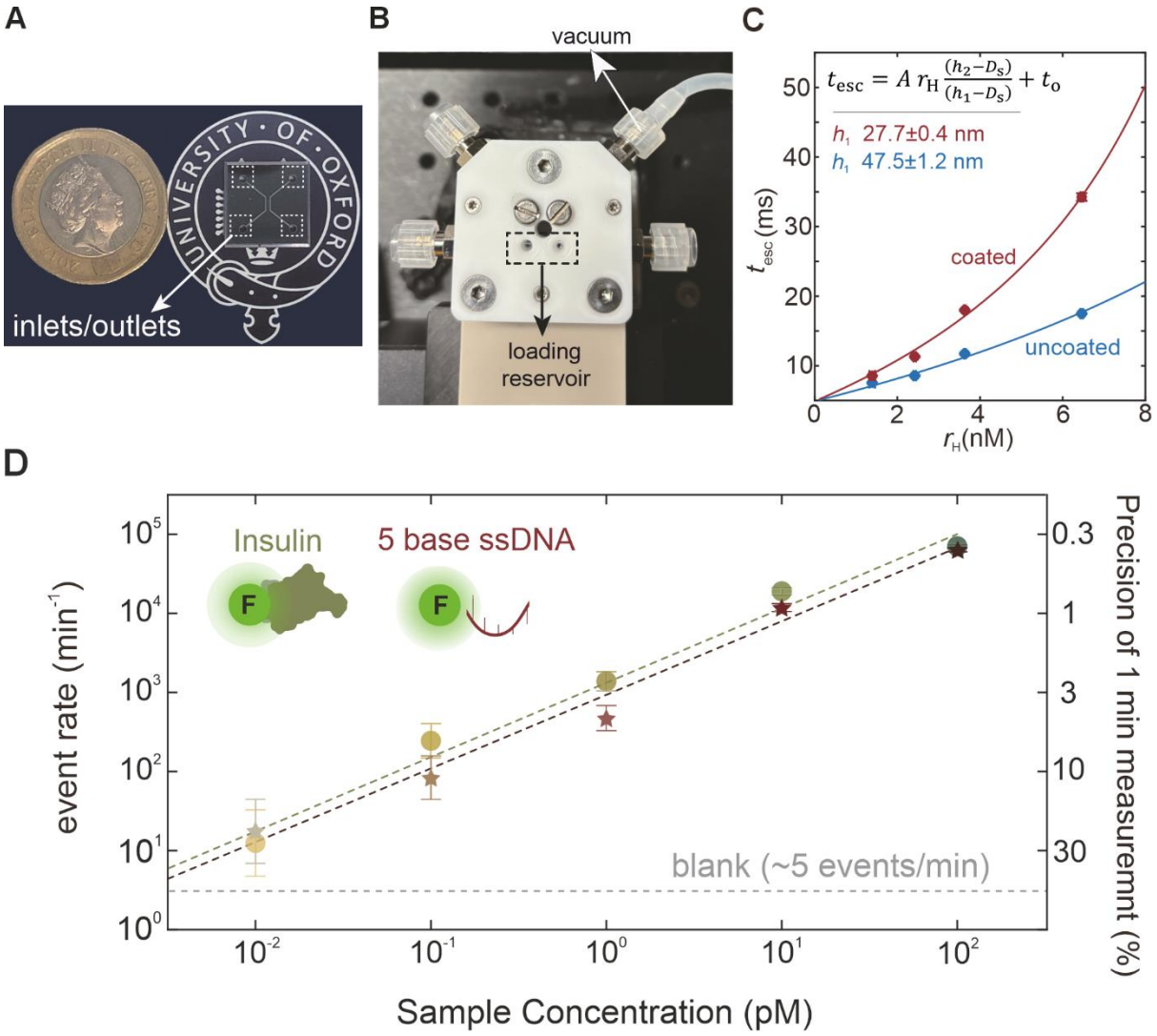
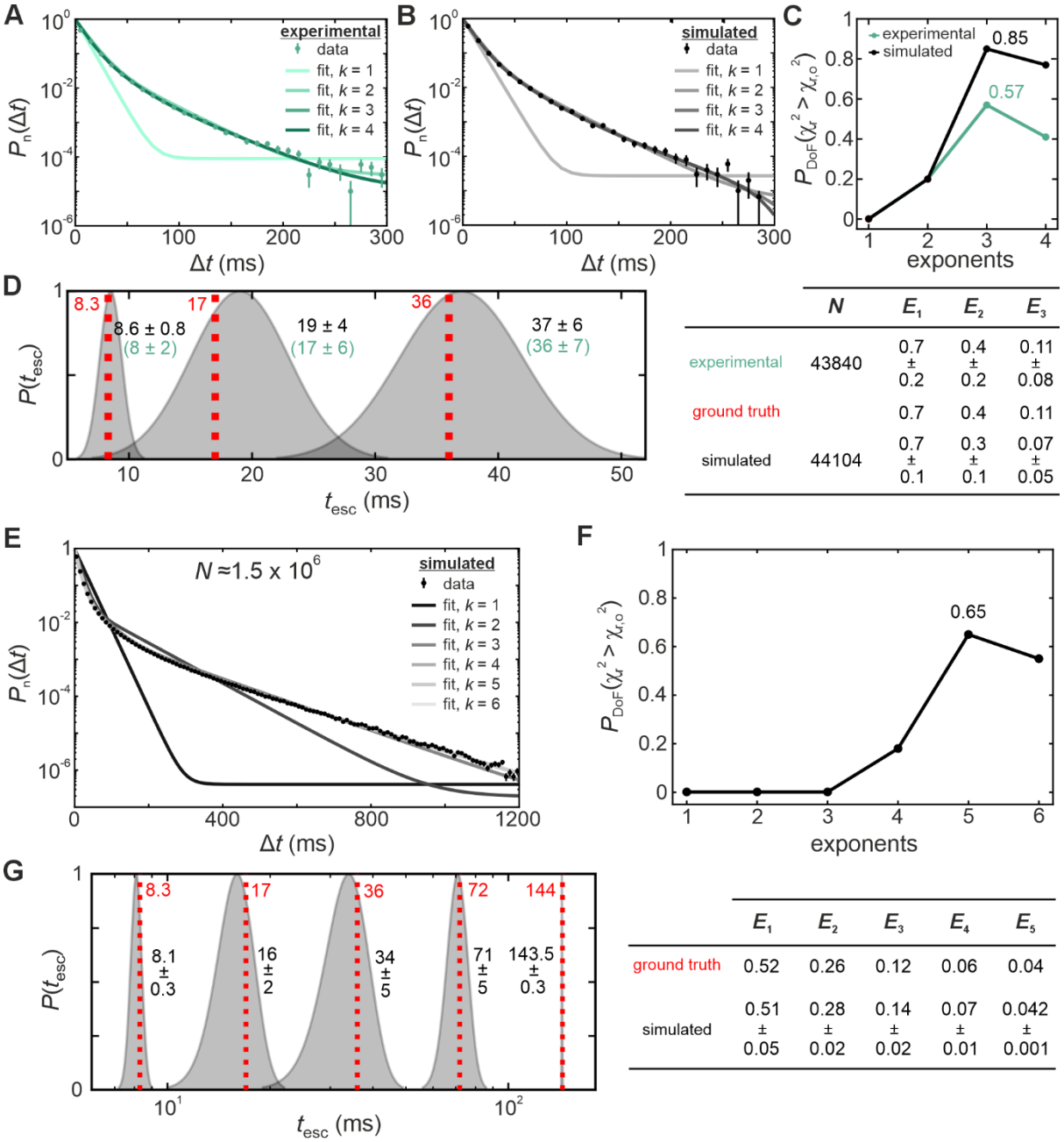


Fig. S5. **Passivation of slits with m-PEG-silane and demonstration of low copy-number molecular counting.** **(A)** Microfluidic chip used in this work, next to 1£ coin for scale. **(B)** Teflon chuck housing the chip (not visible) allows for easy mounting on the widefield fluorescent microscope stage, and application of vacuum. Loading reservoirs are open to air while the outlet is sealed and connected to vacuum (~50 mbar) to facilitate loading of sample into the chip. **(C)** Calibration curves for two nominally identical devices, one with m-PEG coating and one without, were constructed using a series of globular proteins (INS, CA, TF, FER) with different r_H values, as calculated using HYDROPRO. Measured escape times were fitted with Eq. (S1) with $D_s = 2r_H$, $t_o = 5.83$ ms and $d = h_2 - h_1 = 260$ nm as measured by AFM. **(D)** Passivation of devices allows quantitative molecular counting down to 10 fM sample concentration. Samples of labelled 5 ssDNA and INS measured over 5 orders of magnitude in concentration. The event rate measured scales linearly with concentration as expected, showing that ETsy devices can be used as precise and sensitive concentration sensors.



	N	E_1	E_2	E_3
experimental	43840	0.7 ± 0.2	0.4 ± 0.2	0.11 ± 0.08
ground truth		0.7	0.4	0.11
simulated	44104	0.7 ± 0.1	0.3 ± 0.1	0.07 ± 0.05

	E_1	E_2	E_3	E_4	E_5
ground truth	0.52	0.26	0.12	0.06	0.04
simulated	0.51 ± 0.05	0.28 ± 0.02	0.14 ± 0.02	0.07 ± 0.01	0.042 ± 0.001

Fig. S6. **Multiexponential histogram fitting and ground-truth validation. (A)** Measured normalized probability density distributions $P_n(\Delta t)$ vs. Δt (green curves) for $N \approx 4.4 \times 10^4$ recorded escape events of SAM-IV riboswitch (green data and lines), also shown in Fig. 4F, were fitted with Eq. (S1) up to $k = 4$ components. **(B)** Fits of $P_n(\Delta t)$ vs. Δt (grey curves) for $N \approx 4.4 \times 10^4$ simulated escape events (black data) using as simulation inputs the fit escape times (t_i) and event fractions (E_i) obtained by fitting a 3-exponential decay to the experimental SAM-IV riboswitch in (**A**). **(C)** Probability of the expected reduced χ-squared being greater than the observed $P_{\mathrm{DoF}}\left(\chi_r^2 > \chi_{r,o}^2\right)$ for the experiment in (**A**) (green) and simulation in (**B**) (grey), confirms that both experimental and simulated data are best fit with Eq. (S1) when the maximum number of components is 3. **(D)** Ground truth values (3 vertical dashed red lines), based on the experimental fit results (green values), used to simulate 3-component histograms with $N \approx 4.4 \times 10^4$. Each grey normal distribution represents the fit escape time (t_i) and fit error for each component from a simulated histogram, illustrating good agreement with the given ground truth values. Table lists the event fractions (E_i) for each case for experiments and simulations. **(E)** $P_n(\Delta t)$ vs. Δt (grey curves) for $N \approx 1.5 \times 10^6$ simulated escape events (black points) based on a ground truth given by the 3-exponent fit experimental escape times (t_i) and event fractions (E_i) from **A** (ground truth) and including two additional components. The data were fitted with Eq. (S1) for up to a maximum of 6 components. **(F)** $P_{\mathrm{DoF}}\left(\chi_r^2 > \chi_{r,o}^2\right)$ for the fit curves in (**E**) shows that the generated simulated data are best fit with Eq. (S1) when $k = 5$ (~0.65). **(G)** Ground truth timescales displayed as 5 red dotted lines. Each grey normal distribution represents the mean fit escape time (t_i) and corresponding fit error, illustrating good agreement with the ground truth values. Table lists the ground truth and mean simulated event fractions (E_i) with their standard deviations obtained from 5 independent realisations of simulated histograms corresponding to $k = 5$ curve in **E.** Note that as expected, this standard deviation on the mean value is essentially the same as the error on individual fits which give the standard error of the mean.

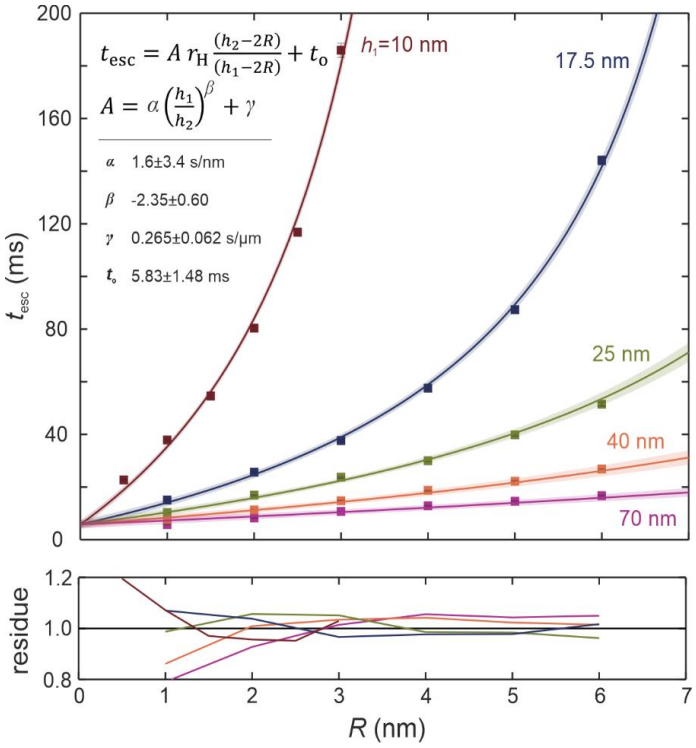


Fig. S7. **BD simulations of spherical particles diffusing in a nanostructured landscape.** BD simulations of spheres of different radii, $R = r_H$, diffusing in a landscape of traps are analysed to determine t_{esc}. $N = 10{,}000$ escape events were simulated for a range of different geometries given by $10 < h_1 < 70$ nm and $d = h_2 - h_1 = 300$ nm. A global fit (lines, with shaded region displaying 1 standard deviation of uncertainty on the overall fit), using Eq. (S21) with $D_s = 2R$, to all five datasets yielded good agreement with the theoretical equation; the average r.m.s. of the residuals is less than 3% (bottom panel). Fit parameters were obtained as follows: $\alpha = 0.16 \pm 0.34$ s/nm, $\beta = -2.35 \pm 0.60$ and $\gamma = 0.265 \pm 0.062$ s/μm, $t_o = 5.83 \pm 1.48$ ms. For further analysis, only the mean parameter values were used to specify Eq. (S21) in any given device geometry.

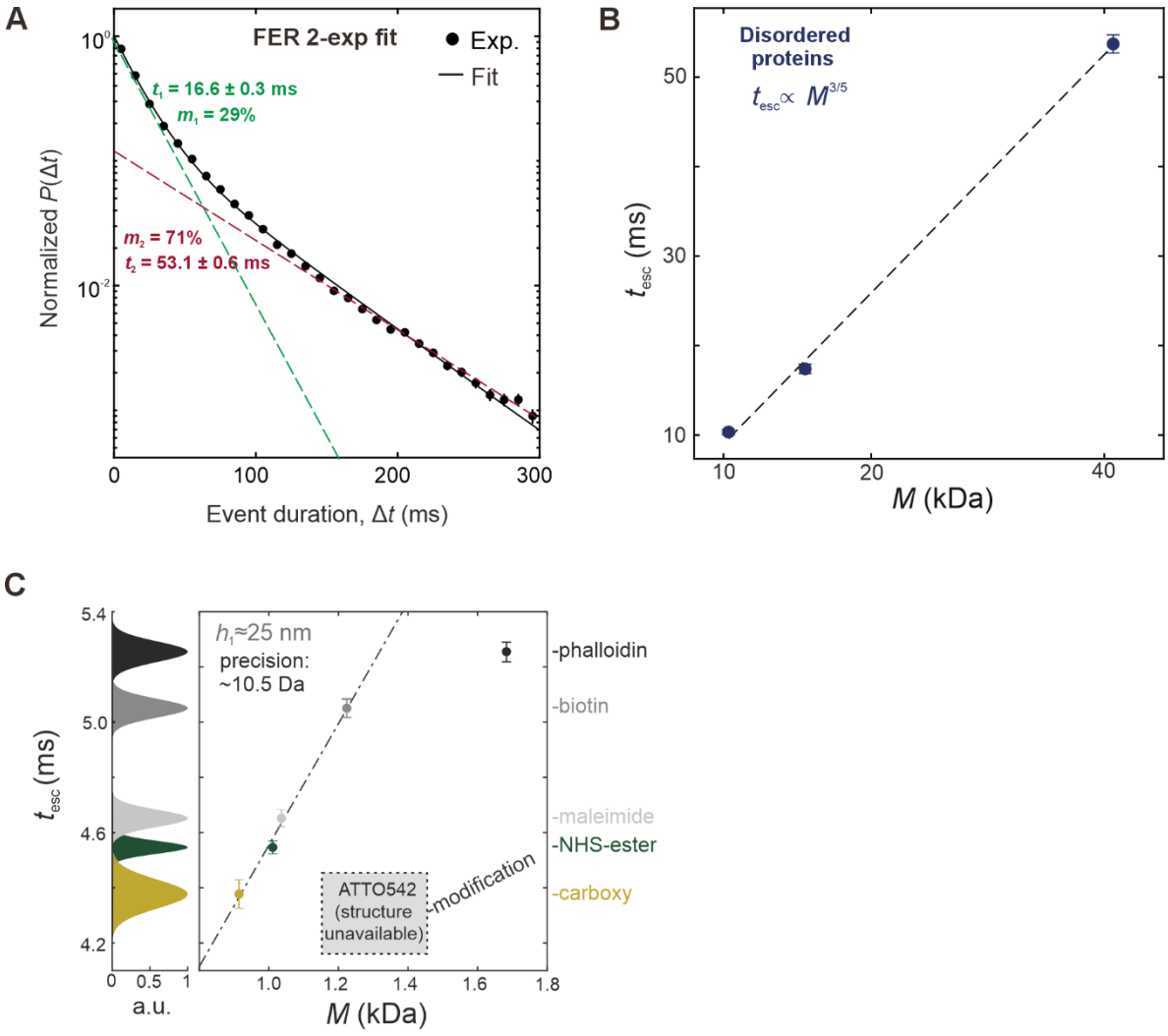


Fig. S8. Measurements on Apoferritin (FER), disordered proteins and ATTO542 dye derivatives. (A) Bi-exponential fit (black curve) of Apo-ferritin escape time histogram (data points). t_1 (green) is as expected for an Apo-ferritin monomer ($M \approx 20$ kDa), while t_2 (red) is taken to correspond to the 24-mer complex ($M \approx 465$ kDa). **(B)** Escape times for disordered proteins (13-mer proline rich polypeptide, ProTα and Stml-1) measured at 5-20 nM concentration in PBS. A linear fit (dashed line) of t_{esc} vs. $M^{\frac{3}{5}}$ agrees with the data. A ~5 min movie was recorded for each sample. **(C)** Plot of t_{esc} vs. M for small fluorescent organic molecules represented by various chemical derivatives of the fluorophore ATTO542. As in Fig. 2B from the main text, ~0.5% measurement precision on t_{esc} implies the ability to resolve differences ca. 10 Da in the $M = 1$ kDa molecular weight regime.

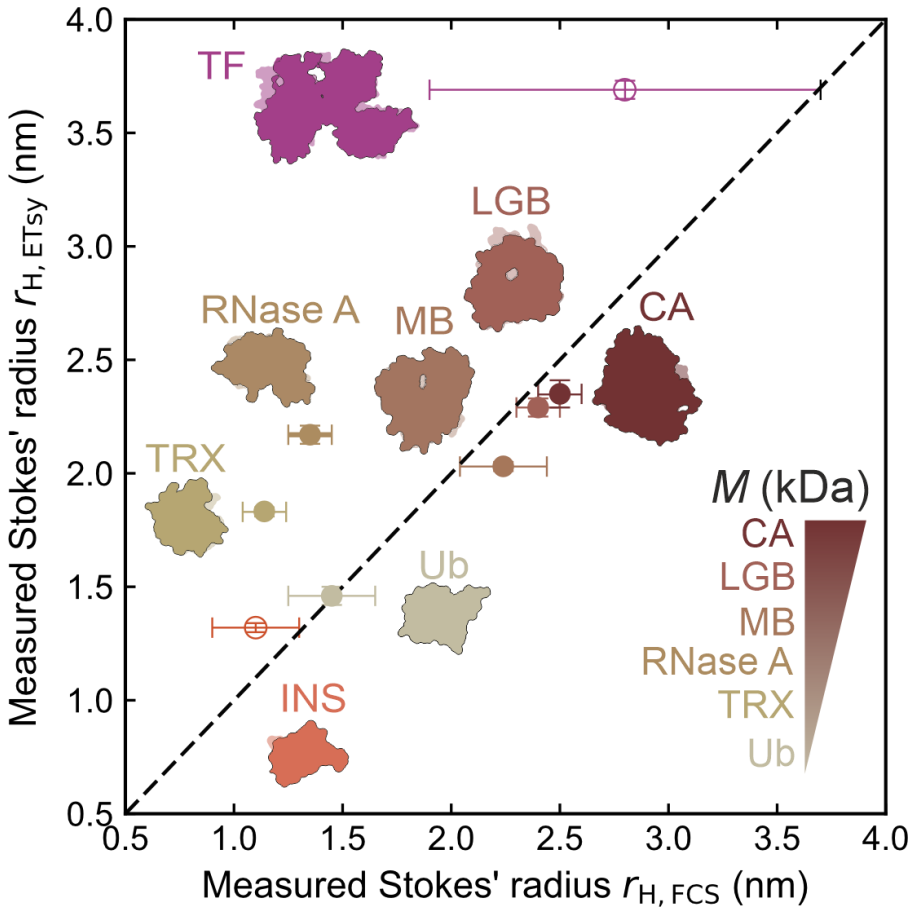


Fig. S9. Comparison of Stokes' radii measured using ETsy and 2f-FCS for a set of globular proteins (Fig. 2D, main text). $r_{H,ETsy}$ for the test set of globular proteins in the range $M = 10 - 30$ kDa (INS, Ub, TRX, RNase A, MB, LGB, CA, TF, see main text Fig. 2D) compare reasonably with measurements using 2f-FCS, $r_{H,FCS}$. The grey diagonal dashed line illustrates perfect agreement between the two techniques. Open symbols represent molecules used for device calibration. 2f-FCS measurements are in general prone to greater inaccuracies on account of very long measurement times needed, and the presence of free dye and multimers in solution – all of which are readily addressed and handled using ETsy.

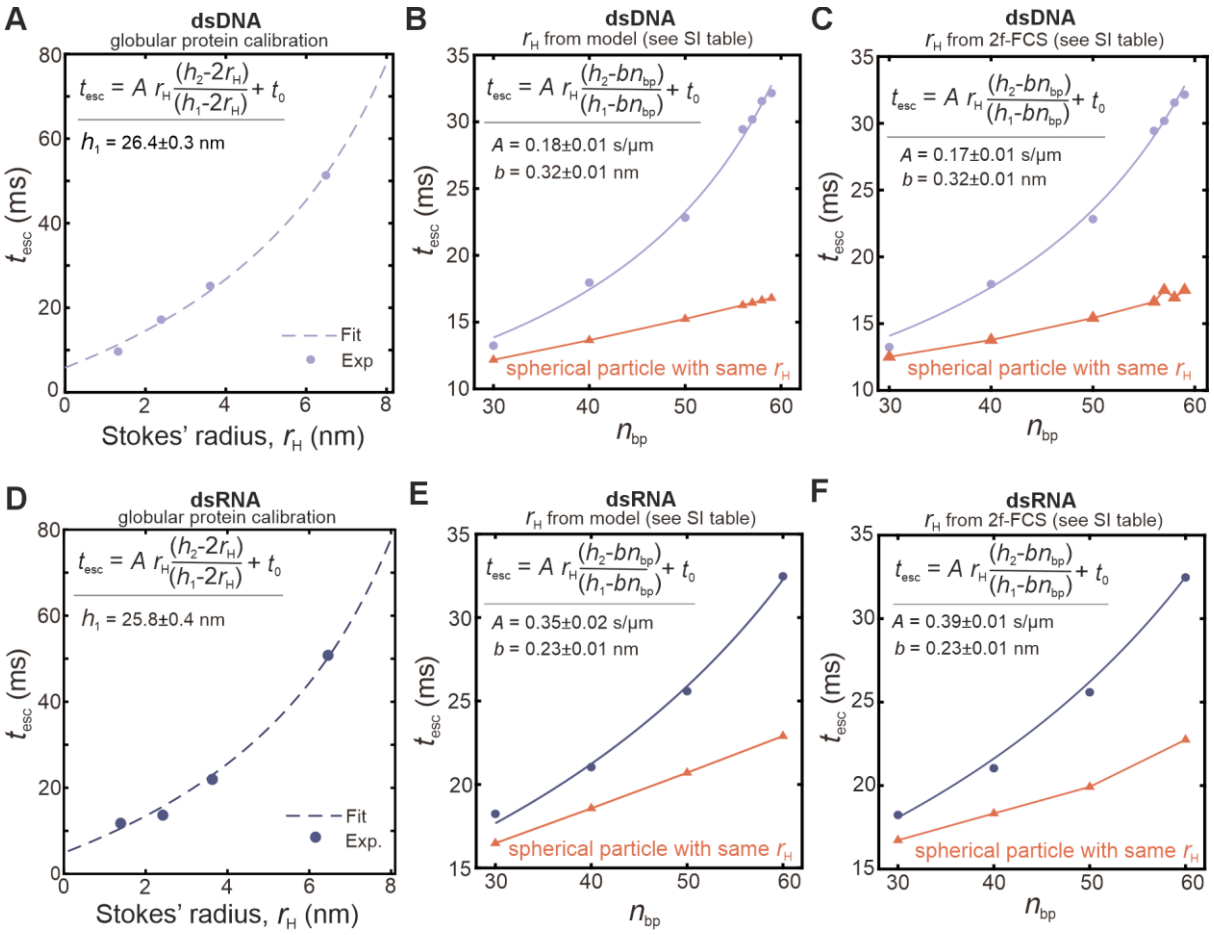
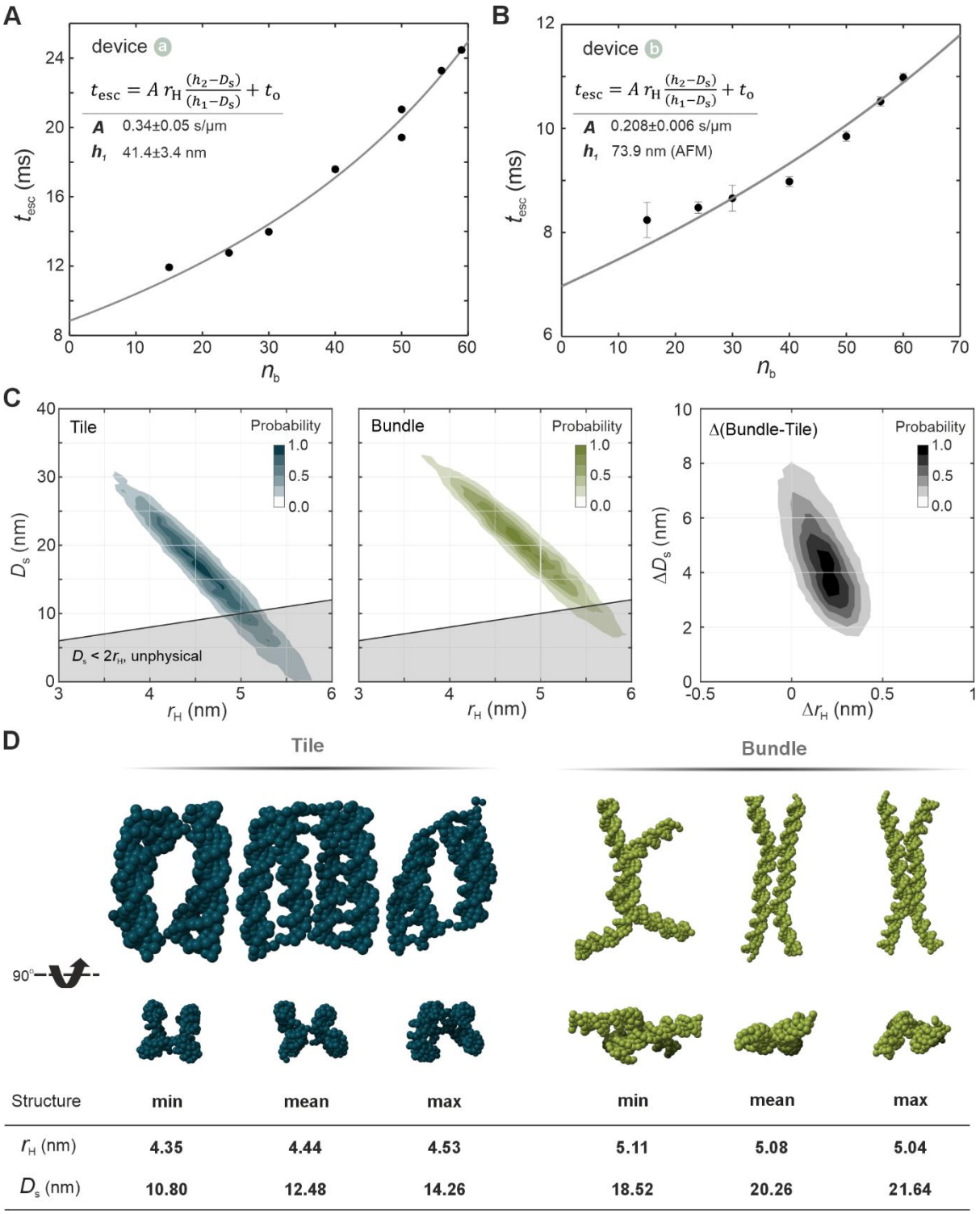


Fig. S10. **Measurements of the rise-per basepair, *b*, in dsDNA and dsRNA.** Calibration curves (dashed lines) for two devices used for *b* measurements of dsDNA **(A)** and dsRNA **(D)**: a series of globular proteins (INS, CA, TF, FER) with different r_H values, as determined from molecular structures using HYDROPRO, were measured and escape times fitted with Eq. (S21) to determine h_1, using $D_s = 2r_H$, $t_o = 5.83$ ms and $d = h_2 - h_1 = 286$ nm and 250 nm, respectively, as measured by AFM. **(B-F)** Measurements of escape time t_{esc} vs. number of basepairs, n_{bp}, for a series of dsDNA **(B, C)** and dsRNA **(E, F)**. Data is fitted to determine A and b using Eq. (S21), with $t_o = 5.83$ ms, $D_s = bn_{bp}$, h_1 and d from calibration (panels **A** and **D**) and using r_H values either as calculated using the theoretical equation (Eq. (S14), panels **B** and **E**) or as measured using 2f-FCS (Table S5, panels **C** and **F**). Whilst the fit value A changes slightly depending on the source of r_H, the inferred b value is robust to this choice. In all cases, orange data illustrates the approximately linear response expected from Eq. (S21) for a spherical particle ($D_s = 2r_H$) with the same r_H as for the corresponding length of the nucleic acid.



Structure	min	mean	max	min	mean	max
r_H (nm)	4.35	4.44	4.53	5.11	5.08	5.04
D_s (nm)	10.80	12.48	14.26	18.52	20.26	21.64

Fig. S11. **Measurement of D_s and r_H for DNA nanostructures.** Calibration curves for two devices (a, b) used for DNA-nanostructure measurements: a series of dsDNA with different number of base pairs, n_{bp}, was measured and escape times fitted with Eq. (S21) where $D_s = 0.34 n_{bp}$ $t_o = 5.83$ ms, and $d = h_2 - h_1 = 286$ nm. **(A)** For device (a), A and h_1 were determined to be 0.34±0.05 s/μm and 41.4±3.4 nm, respectively. **(B)** For device (b), A was determined to be 0.208±0.006 s/μm and h_1 was set to the nominal value as determined by AFM, 73.9 nm. **(C)** Full range of D_s, r_H values obtained for the nanostructures accounting for the calibration uncertainties. 10,000 sets of $h_{1,a}, h_{1,b}$ and A_a, A_b pairs were selected from normal distributions following their respective uncertainties from calibration shown in (**A, B**). For each set, D_s, r_H were calculated for every sample based on t_{esc} values similarly selected from normal distributions following the measurement uncertainties. Then a histogram was constructed to create the probability density maps in the left and middle panels. The difference between D_s and r_H values for the two species was also calculated and similarly plotted in the right panel. Our definition of D_s mandates that for non-spherical objects $D_s > 2r_H$. This excludes (D_s, r_H) solutions in the parameter space indicated by the grey shaded region. **(D)** A series of snapshots from oxDNA simulations of the nanostructures presented in order of decreasing compaction (based on D_s value). r_H values were calculated for each structure using HYDROPRO, while D_s was evaluated using a custom Python script ("miniball").

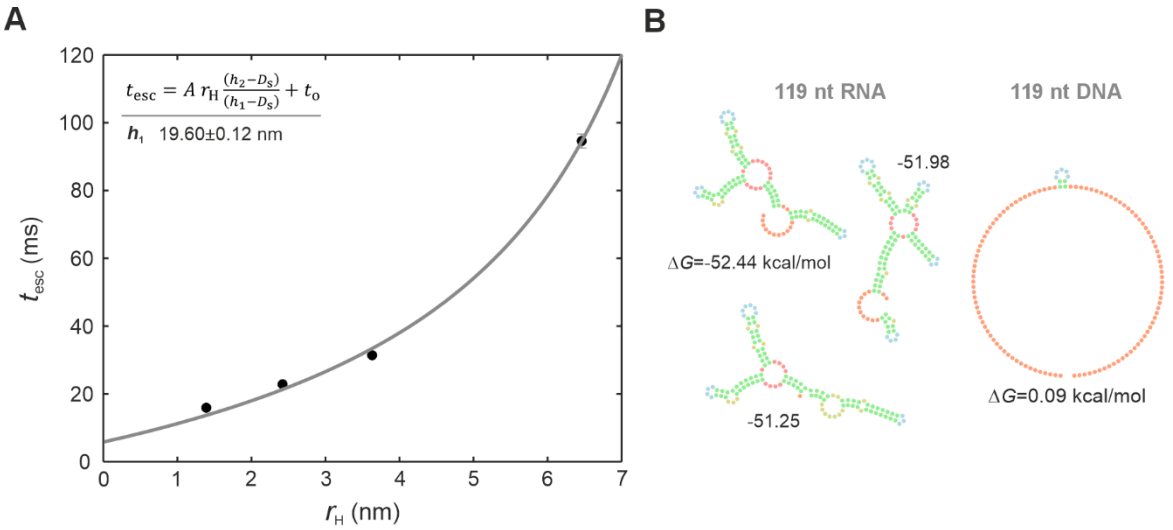


Fig. S12. **Measurement of the structural properties of the SAM-IV riboswitch**. **(A)** Calibration curve for the device used to measure the SAM-IV riboswitch: a series of globular proteins (INS, CA, TF, FER) with different r_H values, as determined using HYDROPRO, were measured and escape times fitted with Eq. (S21) where $D_s = 2r_H$, $t_o = 5.83$ ms and $d = h_2 - h_1 = 259$ nm as measured by AFM. h_1 was determined to be 19.60 ± 0.12 nm. **(B)** The most stable secondary structures predicted by Mfold for the SAM-IV riboswitch and reference sample of 119 nt DNA, and their respective thermodynamic stabilities, ΔG, for 1 M [Na^+] at 37°C. Multiple stable secondary structures were observed for the RNA, compared to essentially no secondary structures obtained for the selected DNA sequence. Note that in solution RNA folds into tertiary structures.

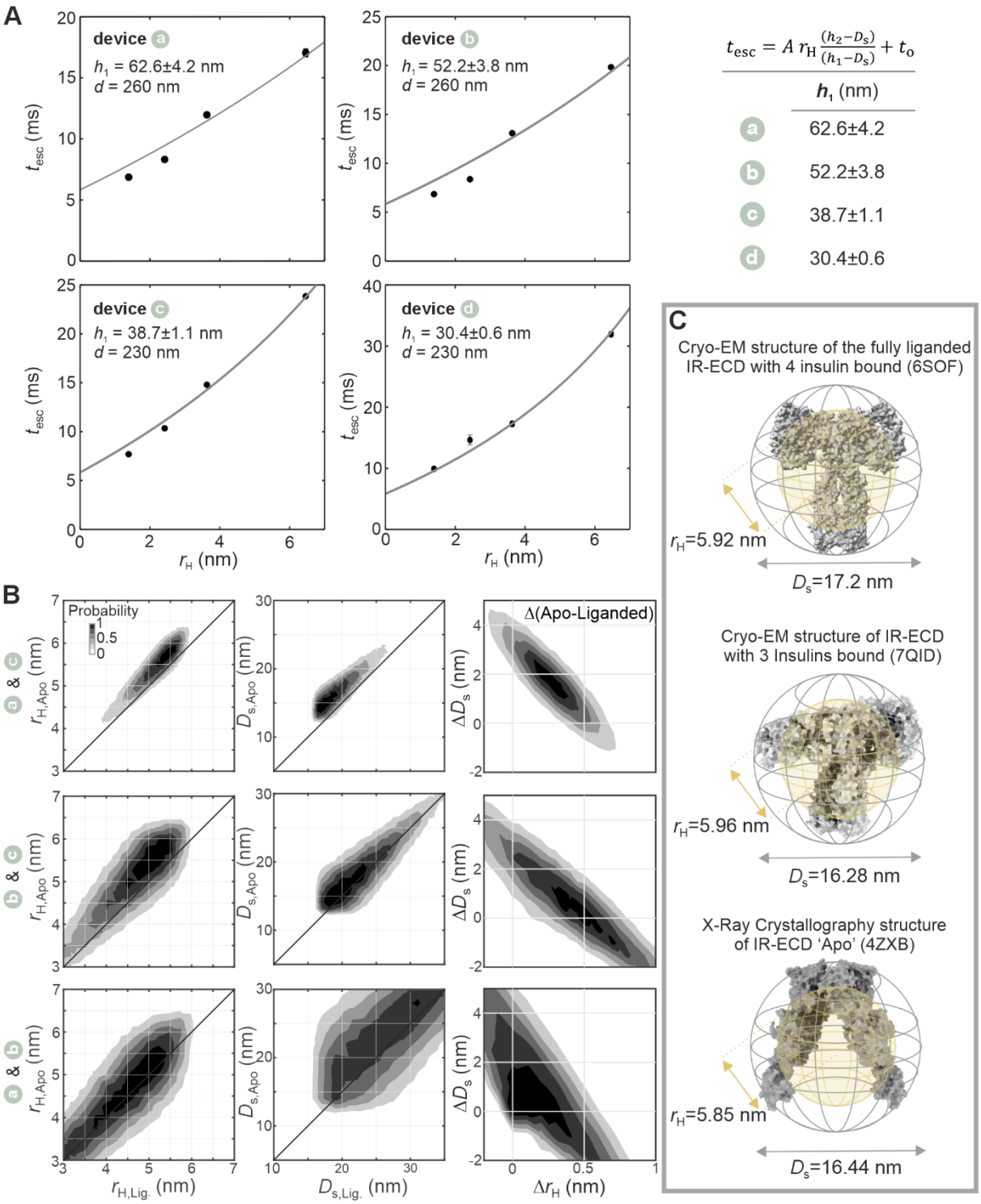


Fig. S13. **Measurement of D_s and r_H of the IR-ECD. (A)** Calibration curves for the devices used to measure the IR-ECD: a series of globular proteins (INS, CA, TF, FER) with different r_H values, as calculated using HYDROPRO, were measured and escape times fitted with Eq. (S21) with $D_s = 2r_H$, $t_o = 5.83$ ms and $d = h_2 - h_1$ as measured by AFM. Fitted h_1 values are displayed in the figure legends. **(B)** Pairwise determination of (D_s, r_H) values obtained for the IR-ECD for each device pair (a, c), (b, c) and (a, b) accounting for the calibration uncertainties. For each pair of devices (i, j), 10,000 sets of $h_{1,i}, h_{1,j}$ pairs were selected from normal distributions following their respective uncertainties from calibration in (**A**). For each set, (D_s, r_H) was calculated for molecular state (Apo. and Lig.) based on measured t_{esc} values similarly selected from normal distributions following the measurement uncertainties. A histogram was constructed to create the probability density maps in the left and middle panels. The difference between the D_s and r_H values for the two states was also calculated and similarly plotted in the right panel. Note that we only include data which satisfies our physical constraint $D_s > 2r_H$. Regardless of the pair of device geometries selected, we consistently note $D_{s,Apo} - D_{s,Lig.} \approx 1.5$ nm and $r_{H,Apo} - r_{H,Lig.} \approx 0.25$ nm. The pair of geometries (a) and (c) yields the most precise readout of the three geometry-pairs examined as explained in the text. **(C)** Calculated D_s and r_H values for structures reported in literature. From top to bottom: fully liganded (PDB: 6SOF) (*61*), partially liganded (PDB:7QID) (*61*), Apo (PDB:4ZXB) (*65*).

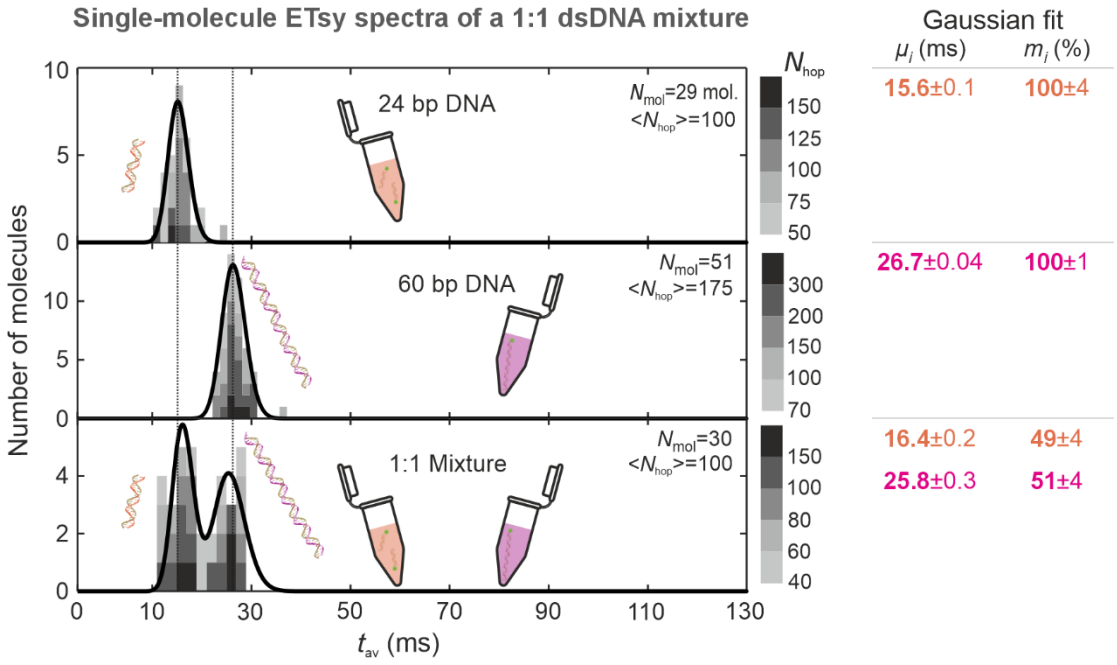


Fig. S14. **Validation of single-molecule spectra measurements using dsDNA.** Single-molecule spectra of t_{av} measurements on individual molecules are constructed for measurements of 24 bp DNA, 60 bp DNA, and a 1:1 mixture of the two. These spectra are fitted with a sum of Erlang distributions (Eq. (S24)) with $k = \langle N_{\mathrm{hop}} \rangle$, which are well approximated by Gaussian distributions for $N_{\mathrm{hop}} \geq 50$ achieved in this study. Fitted parameters for the means, μ_i, and molecule fractions, m_i, are displayed in the table on the right.

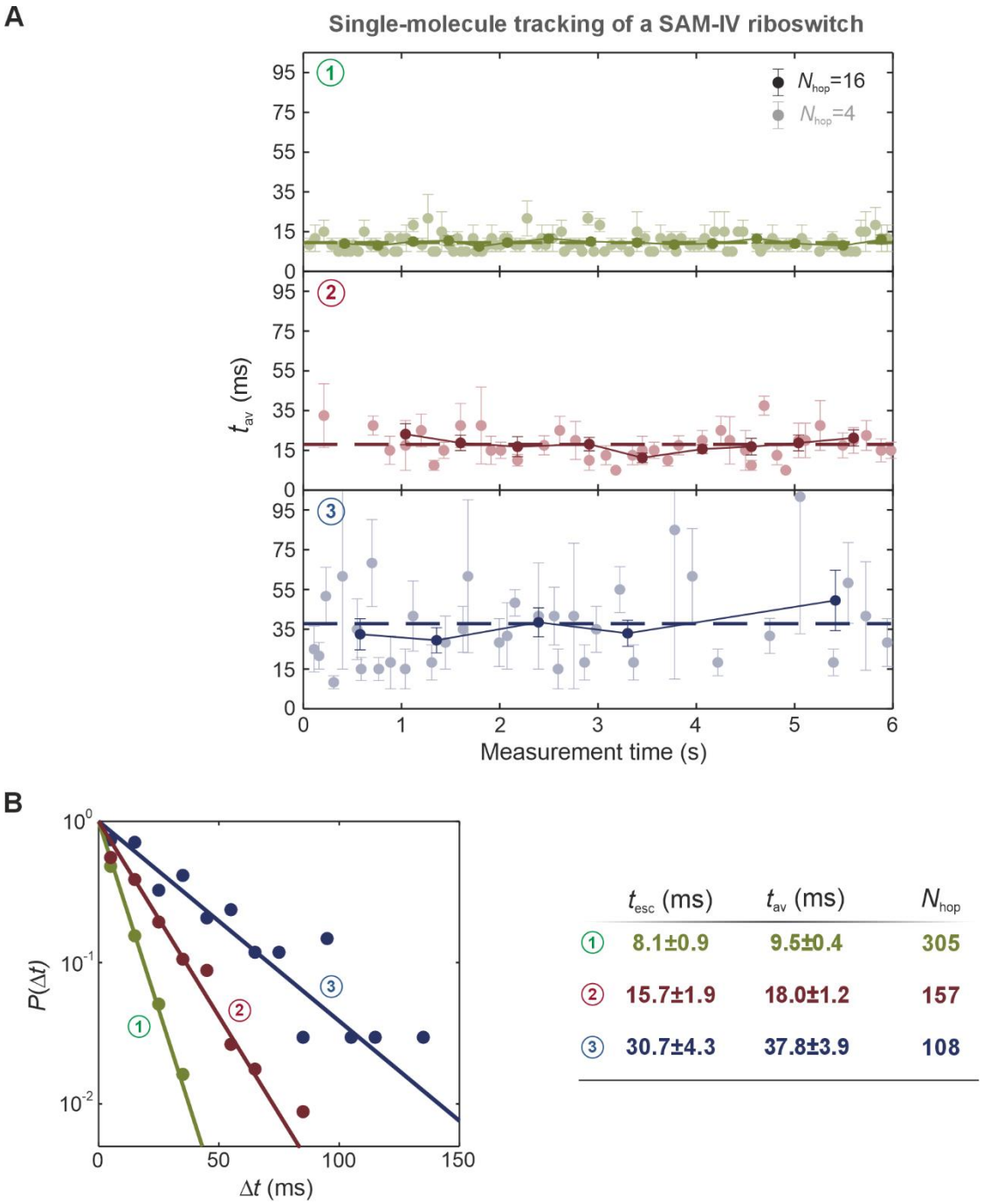


Fig. S15. **Single-molecule tracking of the SAM-IV riboswitch. (A)** Recorded time traces of sequential escape events Δt, averaged over 4 (light symbols) and 16 (dark symbols) consecutive escape events, for three individual molecules in the SAM-IV riboswitch sample, measured in PBS containing 5 mM $MgCl_2$. We observe statistically limited fluctuation in average escape times over the 6 s of observation time suggesting no detectable conformational fluctuation, or interconversion between the different molecular states. **(B)** Escape time histograms are also constructed from the same single molecule data and fit with a single exponential distribution (left panel), yielding $t_{\mathrm{esc}} \approx t_{\mathrm{av}}$ for each molecule as expected (right panel).

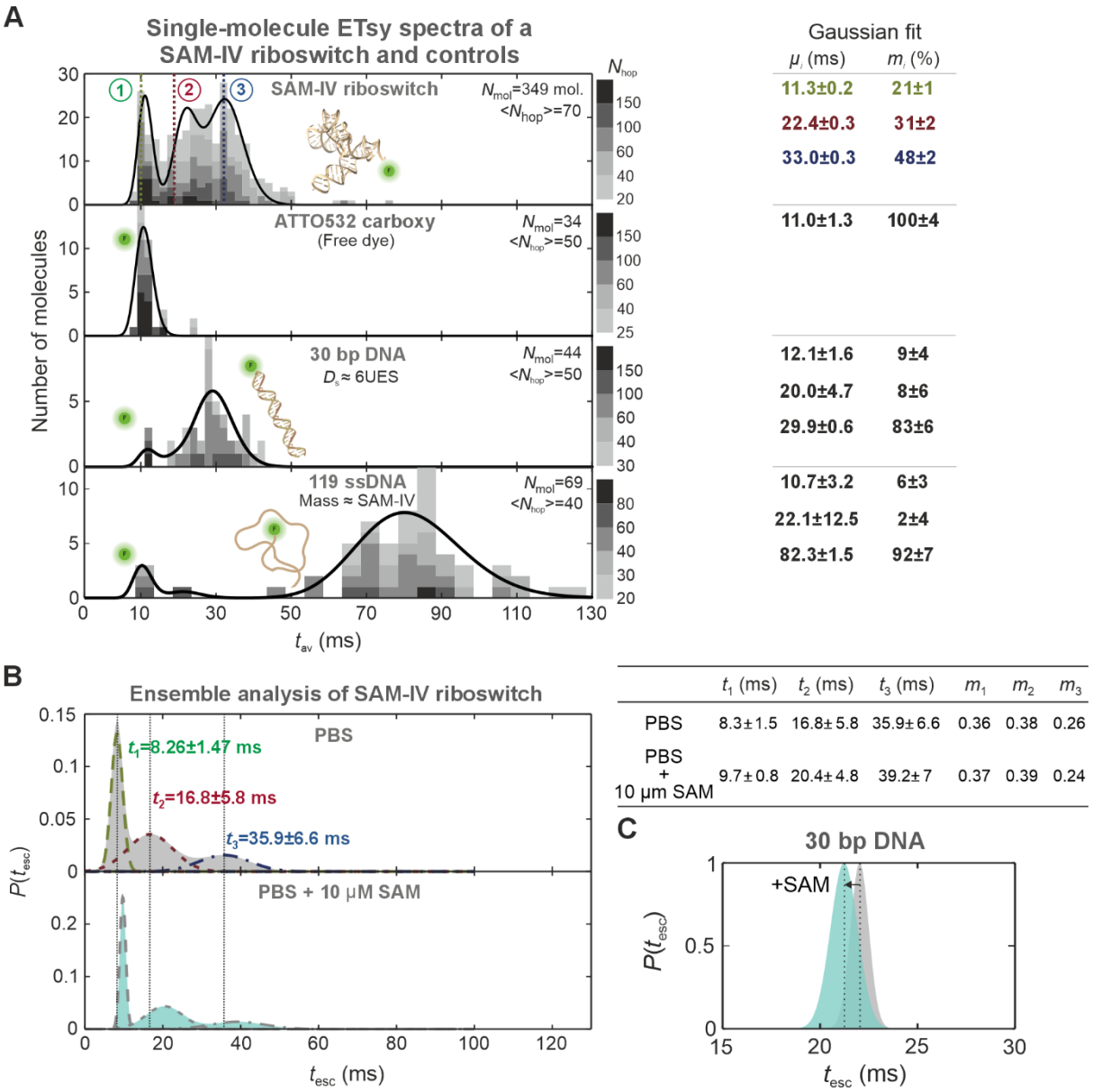


Fig. S16. **Single-molecule measurement of the SAM-IV riboswitch and binding of S-adenosyl methionine (SAM) to the riboswitch. (A)** Single molecule spectra of t_{av} are constructed for measurements of a SAM-IV riboswitch, ATTO532 carboxy, 30 bp DNA and 119 ssDNA. These spectra are fitted with a sum of Erlang distributions (Eq. (S24)) with $k = \langle N_{\mathrm{hop}} \rangle$, which approximate to Gaussian distributions for the large N_{hop} values in this study. Fitted parameters for the means, μ_i, and molecule fractions, m_i, are displayed in the table on the right. **(B)** Ensemble analysis of measurements of SAM-IV riboswitch incubated in PBS with and without 10 µM SAM. Escape time histograms were fitted with tri-exponential distributions (fit values in table, right panel) and the obtained timescales are plotted as a probability density plot. We do not observe significant shifts in the timescales t_2 and t_3, characteristic of the riboswitch, upon the addition of SAM. **(C)** As a control experiment, 30 bp DNA was also compared in the different buffers, displaying, as expected, the negligible change in t_{esc} with the addition of SAM.

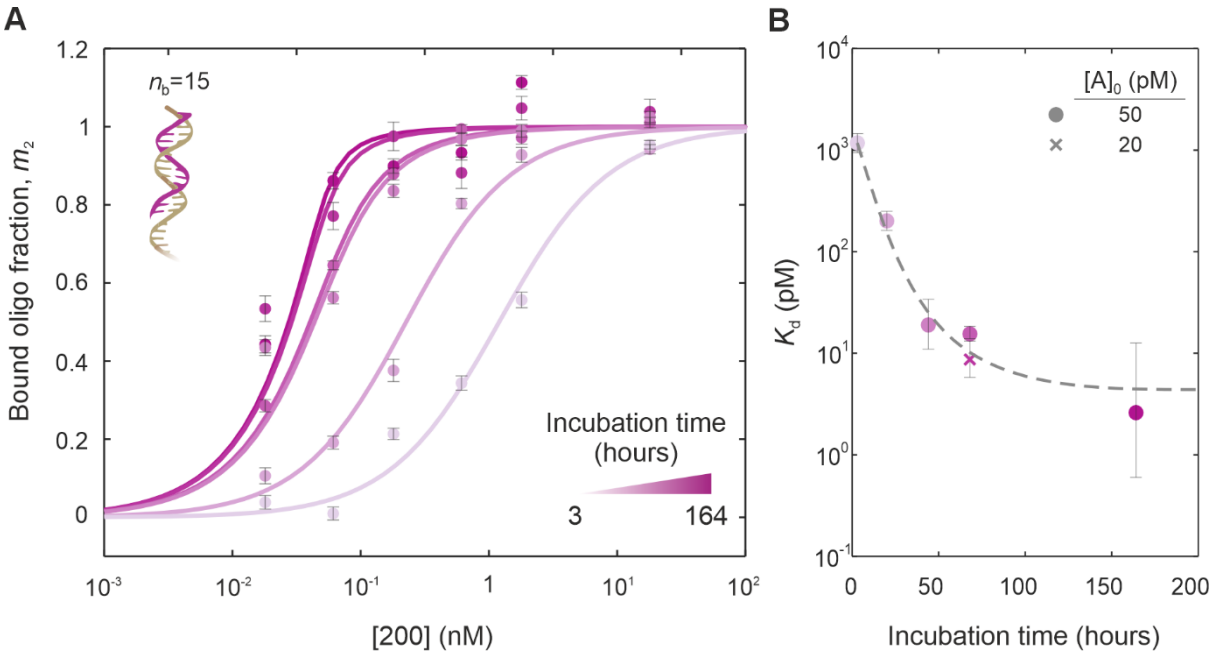


Fig. S17. **Assessment of the attainment of equilibrium in the measurement of K_d for a 15 base ssDNA oligo binding to 200 base ssDNA. (A)** Bound-fraction titration curves for a labelled sample of 15 base ssDNA incubated with complementary 200 base ssDNA for varying amounts of time at room temperature before measurement. In each case, the m_2 values were determined from bi-exponential fitting and fit with Eq. (S29) to determine the K_d (panel **B**). We approach saturation of the measured K_d after ~1 week. However, since in this case, the final inferred $K_d \approx 2.5$ pM is not accurate as it is much smaller than $[A]_0 = 50$ pM, the initial concentration of 15 ssDNA added. In Fig. 5 from the main text, we report the value obtained after 68 hours of incubation with $[A]_0 = 20$ pM (cross symbol). Dashed line in panel **B** is a guide to the eye.

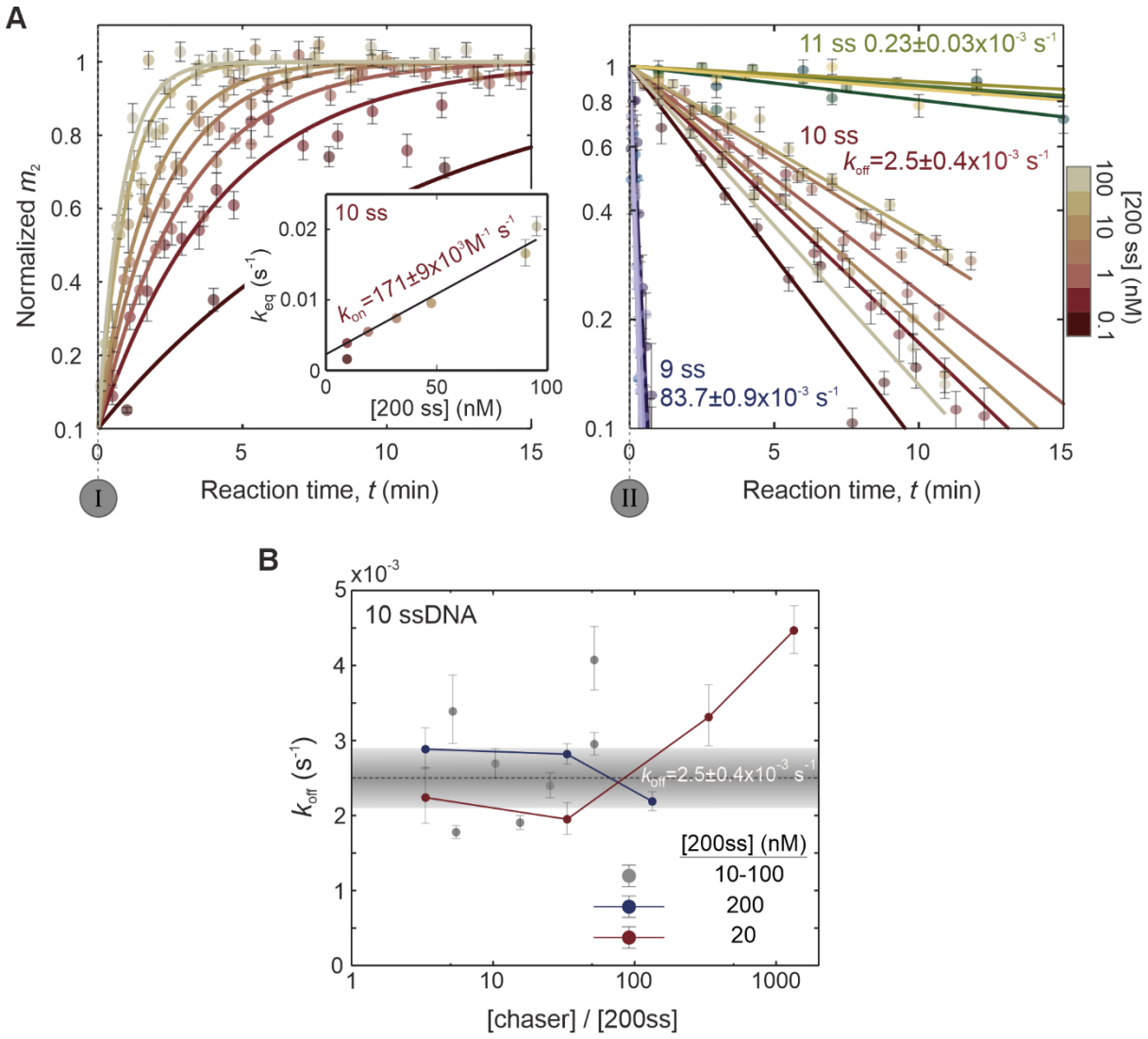


Fig. S18. **Measurement of on and off-rates and assessment of the role of cooperative unbinding in k_{off} measurements. (A)** Left - Measurements of m_2 vs. reaction time, t, performed with 0.5 nM of labelled 10 ssDNA oligo with increasing concentrations of complementary 200 ssDNA are fit with Eq. (S30) to obtain the equilibrium rate constant k_{eq}. Inset shows that k_{eq} increases proportionately with concentration of 200 ssDNA. Right – Equilibrated, fully bound mixtures of 0.5-2 nM 9-11 ss oligo and 200 ssDNAs at various concentrations are mixed with 0.5-5 μM of unlabelled 11 ssDNA oligo, and the decaying bound fraction, m_2, measured as a function of time to yield the cooperative rate constant of dissociation, k_{off}, from a fit to the data with Eq. (S32). k_{off} can vary by more than an order of magnitude for single-base differences. k_{eq} and k_{off} together determine the rate constant of association, k_{on} (left inset), from which K_d can be determined independently. **(B)** k_{off} was measured for 10 ssDNA equilibrated with complementary 200 ssDNA using a variety of different chaser (unlabelled 11 ssDNA) concentrations. Grey datapoints show the data as shown in (**A**), with mean value and standard error on the mean displayed by the dashed line and shaded region. Here, $[A]_0 = 0.5$ nM, $[200ss] = 10 - 100$nM and $[chaser] = 500$ nM. Red and blue datapoints show k_{off} measurements with $[A]_0 = 1$ nM, $[200ss] = 20$ and 200 nM, respectively, and $[chaser] = 80 - 24{,}000$ nM. Varying the ratio of $[chaser]/[200ss]$ has negligible effect on k_{off}, until $[chaser]/[200ss] \geq 100$, which is outside the range of the data used to determine the k_{off} in this work.

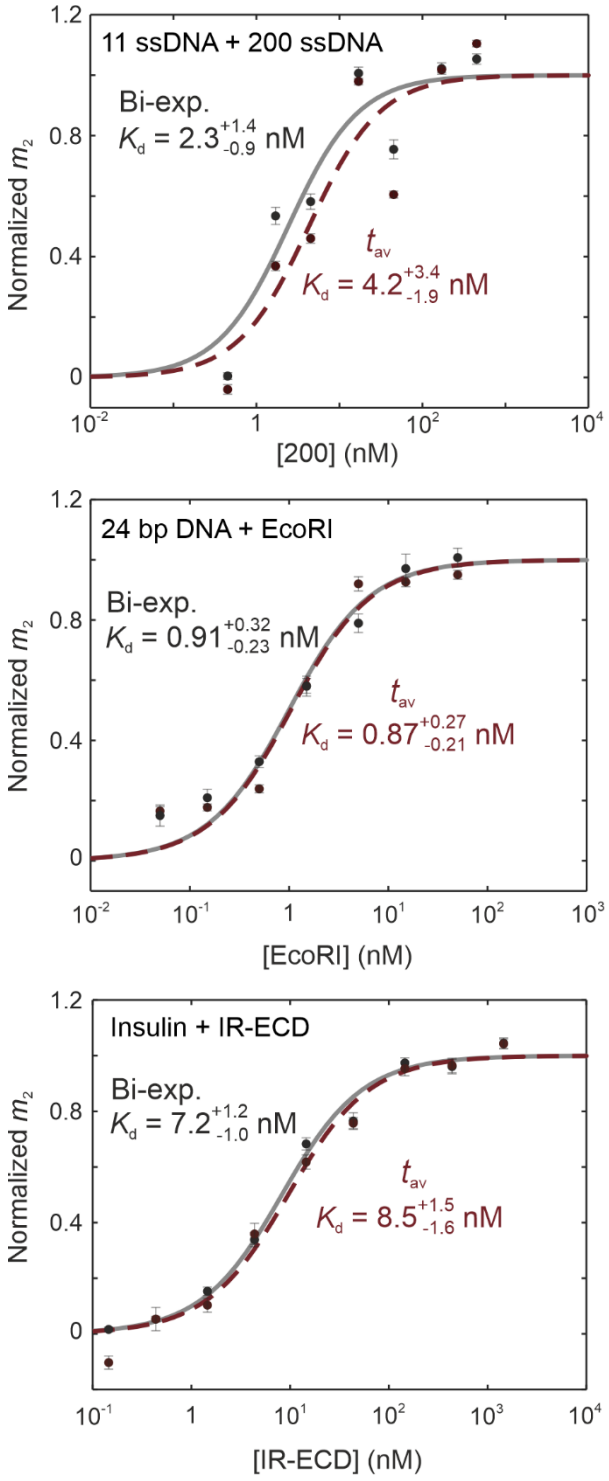


Fig. S19. **Comparison of K_d measurements using bi-exponential and t_{av} fitting methods to determine m_2.** Identical datasets of affinity measurements for three different systems (top: 11 ssDNA + 200 ssDNA, middle: 24 bp DNA + EcoRI, bottom: Insulin + IR-ECD) were analysed using both methods to extract m_2. When fitted with Eq. (S29), we obtain K_d values that are consistent within error when extracting m_2 using bi-exponential fitting (grey, solid curve), and t_{av} (red, dashed curve), for all three systems. This validates the method for use on systems where bi-exponential fitting is challenging.

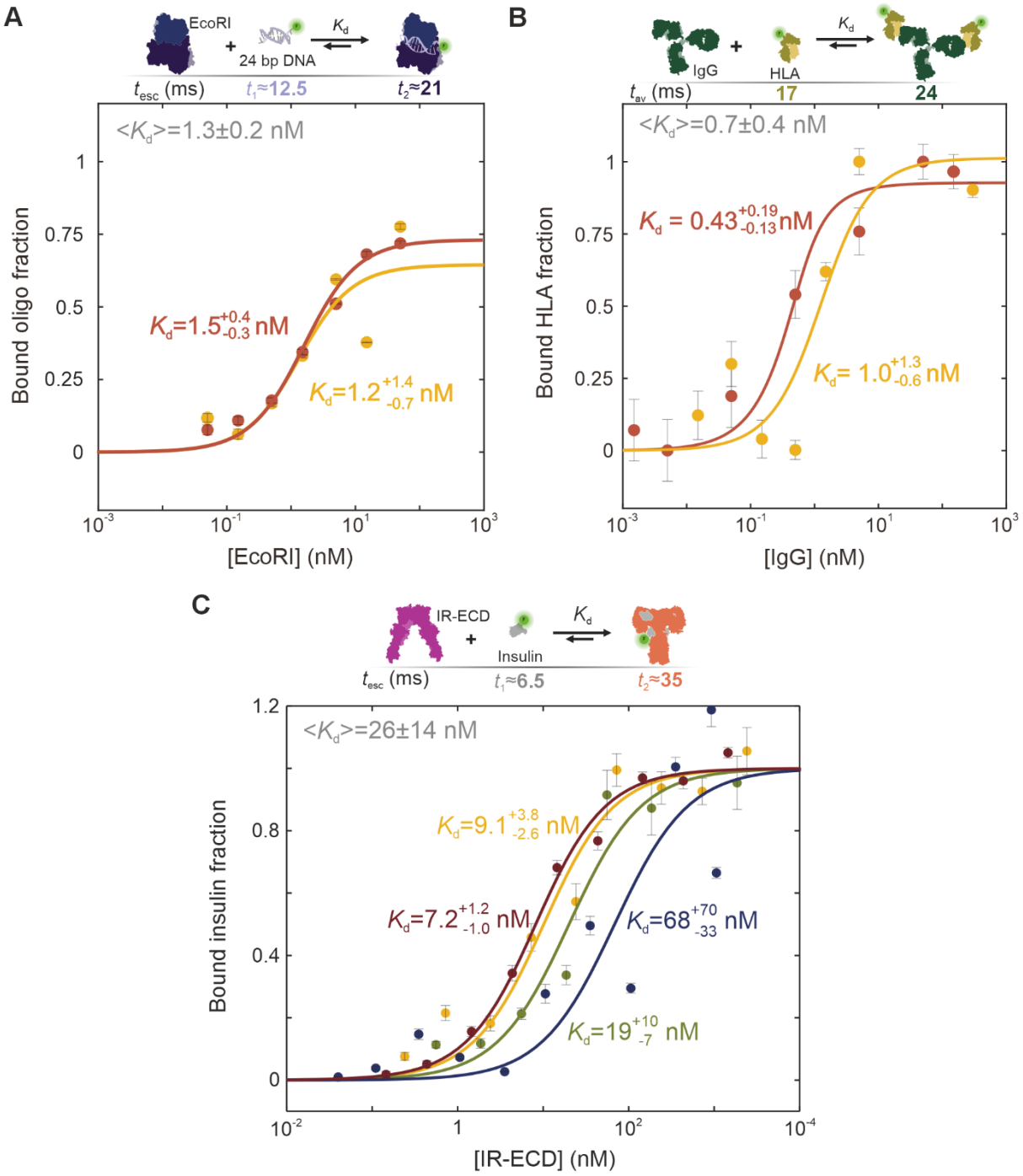


Fig. S20. Measuring K_d for the interactions of EcoRI, HLA-IgG and IR-ECD and their binding partners. Data for 2-4 repetitions of affinity measurements for a labelled sample of **(A)** 24 bp DNA incubated with EcoRI, **(B)** HLA antigen incubated with corresponding IgG and **(C)** insulin incubated with the IR-ECD are shown. In each repetition, 0.25 nM DNA was incubated with varying concentrations of EcoRI in PBS at room temperature for 2 hours before measuring, 1 nM HLA was incubated with varying concentrations of IgG at room temperature for 1 hour before measuring, and 2 nM insulin was incubated with varying concentrations of IR-ECD in HBS at 4°C overnight before measuring. (**A, C**) The m_2 values extracted from bi-exponential fitting were fit with Eq. (S29), and a mean and standard error on the mean of 2 and 4 repeats gives a $K_d = 1.3 \pm 0.2$ nM and $K_d = 26 \pm 14$ nM, respectively. In the main text Fig. 6, we report the lowest K_d values obtained, 1.2 nM and 7.2 nM, respectively. (**B**) The m_2 values extracted from t_{av} were fit with Eq. (S29), and a mean and standard error on the mean of both repeats gives a $K_d = 0.7 \pm 0.4$ nM. In the main text Fig. 6, we report the lowest K_d value obtained, 0.43 nM.

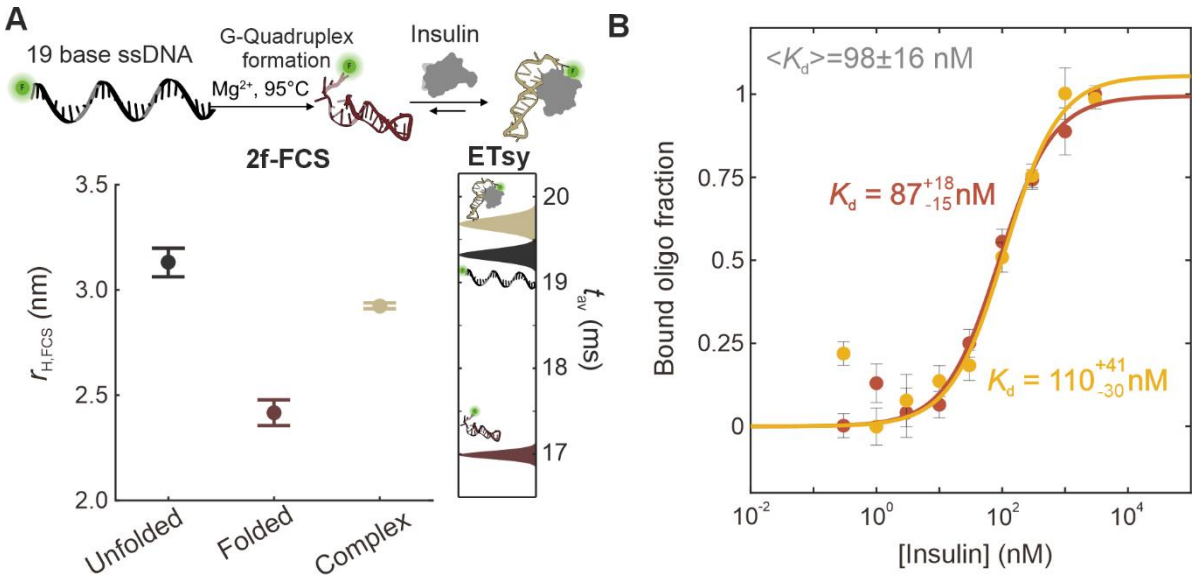


Fig. S21. Measuring K_d for the binding of insulin to an Insulin aptamer. (A) 2f-FCS measurements of the insulin aptamer in unfolded, folded and insulin-bound states displaying a similar general trend as that measured in ETsy (right panel). **(B)** Two repetitions of affinity measurements for a labelled sample of folded aptamer incubated with insulin are shown. In each case, 1 nM of fluorescently labelled aptamer was incubated with varying concentrations of insulin in 10 mM $MgCl_2$ PBS at room temperature for 1 hour before measuring. The m_2 values extracted from t_{av} were fit with Eq. (S29), and a mean and standard error on the mean of both repeats gives a $K_d = 98 \pm 16$ nM. In main text Fig. 6, we report the lowest K_d value obtained, 87 nM.

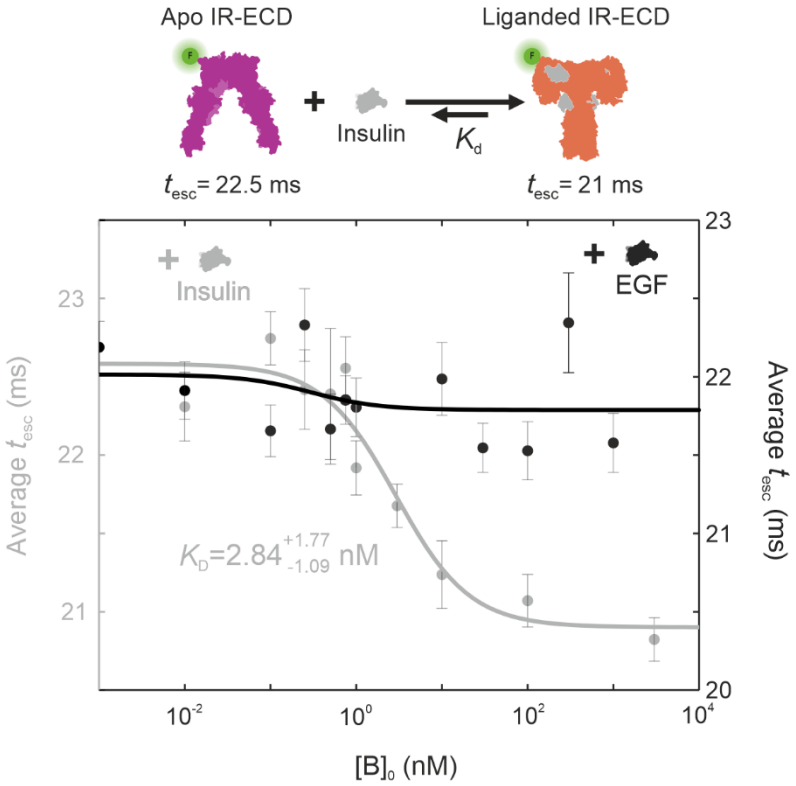


Fig. S22. **Measuring K_d for the interaction of IR-ECD with insulin.** Affinity measurement for the interaction of labelled IR-ECD with insulin (grey data), or epidermal growth factor (EGF) (black data). 10 nM nominal concentration of IR-ECD was incubated with varying concentrations of insulin or EGF in PBS with 0.03% Tween20 at room temperature for 1 hour before measuring. m_2 values were determined from t_{av} and fit using Eq. (S29). We observe negligible change in t_{av} upon addition of EGF showing the specificity of the IR-ECD to insulin.

S10. References

1. Q. Wang, W. E. Moerner, Single-molecule motions enable direct visualization of biomolecular interactions in solution. *Nature Methods* **11**, 555–558 (2014).
2. H. Wilson, Q. Wang, ABEL-FRET: tether-free single-molecule FRET with hydrodynamic profiling. *Nature Methods* **18**, 816–820 (2021).
3. J. A. Jones, D. K. Wilkins, L. J. Smith, C. M. Dobson, Characterisation of protein unfolding by NMR diffusion measurements. *Journal of Biomolecular NMR* **10**, 199–203 (1997).
4. D. K. Wilkins *et al.*, Hydrodynamic Radii of Native and Denatured Proteins Measured by Pulse Field Gradient NMR Techniques. *Biochemistry* **38**, 16424-16431 (1999).
5. R. Russell, I. S. Millett, S. Doniach, D. Herschlag, Small angle X-ray scattering reveals a compact intermediate in RNA folding. *Nature Structural Biology* **7**, 367-370 (2000).
6. J. E. Kohn *et al.*, Random-coil behavior and the dimensions of chemically unfolded proteins. *Proceedings of the National Academy of Sciences* **101**, 12491-12496 (2004).
7. S. B. Dubin, J. H. Lunacek, G. B. Benedek, Observation of the spectrum of light scattered by solutions of biological macromolecules. *Proceedings of the National Academy of Sciences* **57**, 1164 - 1171 (1967).
8. M. Weissman, H. Schindler, G. Feher, Determination of molecular weights by fluctuation spectroscopy: application to DNA. *Proceedings of the National Academy of Sciences* **73**, 2776-2780 (1976).
9. T. Ha *et al.*, Single-molecule fluorescence spectroscopy of enzyme conformational dynamics and cleavage mechanism. *Proceedings of the National Academy of Sciences* **96**, 893-898 (1999).
10. C. B. Müller *et al.*, Precise measurement of diffusion by multi-color dual-focus fluorescence correlation spectroscopy. *Europhysics Letters* **83**, 46001 (2008).
11. G. Young *et al.*, Quantitative mass imaging of single biological macromolecules. *Science* **360**, 423-427 (2018).
12. B. Špačková *et al.*, Label-free nanofluidic scattering microscopy of size and mass of single diffusing molecules and nanoparticles. *Nature Methods* **19**, 751-758 (2022).
13. M. Dahmardeh *et al.*, Self-supervised machine learning pushes the sensitivity limit in label-free detection of single proteins below 10 kDa. *Nature Methods* **20**, 442–447 (2023).
14. L.-M. Needham *et al.*, Label-free detection and profiling of individual solution-phase molecules. *Nature* **629**, 1062–1068 (2024).
15. S. Ohayon *et al.*, Full-Length Single Protein Molecules Tracking and Counting in Thin Silicon Channels. *Advanced Materials* **36**, 2314319 (2024).
16. J. Han, H. G. Craighead, Separation of Long DNA Molecules in a Microfabricated Entropic Trap Array. *Science* **288**, 1026-1029 (2000).
17. J. Fu, R. B. Schoch, A. L. Stevens, S. R. Tannenbaum, J. Han, A patterned anisotropic nanofluidic sieving structure for continuous-flow separation of DNA and proteins. *Nature Nanotechnology* **2**, 121–128 (2007).
18. E. V. Yates *et al.*, Latent analysis of unmodified biomolecules and their complexes in solution with attomole detection sensitivity. *Nature Chemistry* **7**, 802–809 (2015).
19. E. C. Yusko *et al.*, Real-time shape approximation and fingerprinting of single proteins using a nanopore. *Nature Nanotechnology* **12**, 360–367 (2016).
20. F. Ruggeri *et al.*, Single-molecule electrometry. *Nature Nanotechnology* **12**, 488-495 (2017).
21. F. Ruggeri, M. Krishnan, Entropic Trapping of a Singly Charged Molecule in Solution. *Nano Letters* **18**, (2018).

22. S. Schmid, P. Stömmer, H. Dietz, C. Dekker, Nanopore electro-osmotic trap for the label-free study of single proteins and their conformations. *Nature Nanotechnology* **16**, 1244–1250 (2021).
23. G. Krainer *et al.*, Direct digital sensing of protein biomarkers in solution. *Nature Communications* **14**, 653 (2023).
24. R. P. B. Jacquat *et al.*, Single-Molecule Sizing through Nanocavity Confinement. *Nano Letters* **23**, 1629-1636 (2023).
25. I. I. Hosseini *et al.*, Tunable nanofluidic device for digital nucleic acid analysis. *Nanoscale* **16**, 9583-9592 (2024).
26. G. Krainer *et al.*, Single-molecule digital sizing of proteins in solution. *Nature Communications* **15**, 7740 (2024).
27. M. Friedel, A. Baumketner, J.-E. Shea, Effects of surface tethering on protein folding mechanisms. *Proceedings of the National Academy of Sciences* **103**, 8396-8401 (2006).
28. Y. Pang, R. Gordon, Optical Trapping of a Single Protein. *Nano Letters* **12**, 402-406 (2012).
29. F. Perrin, Mouvement Brownien d'un ellipsoide (II). Rotation libre et dépolarisation des fluorescences. Translation et diffusion de molécules ellipsoidales. *Journal de Physique et le Radium* **7**, 1-11 (1936).
30. A. Ortega, D. Amorós, J. García de la Torre, Prediction of Hydrodynamic and Other Solution Properties of Rigid Proteins from Atomic- and Residue-Level Models. *Biophysical Journal* **101**, 892-898 (2011).
31. P. J. Fleming, K. G. Fleming, HullRad: Fast Calculations of Folded and Disordered Protein and Nucleic Acid Hydrodynamic Properties. *Biophysical Journal* **114**, 856-869 (2018).
32. E. Haustein, P. Schwille, Fluorescence Correlation Spectroscopy: Novel Variations of an Established Technique. *Annual Review of Biophysics* **36**, 151-169 (2007).
33. A. G. Kikhney, D. I. Svergun, A practical guide to small angle X-ray scattering (SAXS) of flexible and intrinsically disordered proteins. *FEBS Letters* **589**, 2570-2577 (2015).
34. M. Bespalova, A. Behjatian, N. Karedla, R. Walker-Gibbons, M. Krishnan, Opto-Electrostatic Determination of Nucleic Acid Double-Helix Dimensions and the Structure of the Molecule–Solvent Interface. *Macromolecules* **55**, 6200-6210 (2022).
35. R. Walker-Gibbons, X. Zhu, A. Behjatian, T. J. D. Bennett, M. Krishnan, Sensing the structural and conformational properties of single-stranded nucleic acids using electrometry and molecular simulations. *Scientific Reports* **14**, 20582 (2024).
36. G. Kloes *et al.*, Far-Field Electrostatic Signatures of Macromolecular 3D Conformation. *Nano Letters* **22**, 7834–7840 (2022).
37. P. J. Flory, *Principles of Polymer Chemistry*. (Cornell University Press, 1953).
38. Y. Nojima, K. Iguchi, Y. Suzuki, A. Sato, The pH-Dependent Formation of PEGylated Bovine Lactoferrin by Branched Polyethylene Glycol (PEG)-N-Hydroxysuccinimide (NHS) Active Esters. *Biological and Pharmaceutical Bulletin* **32**, 523-526 (2009).
39. F. Ruggeri, M. Krishnan, Lattice diffusion of a single molecule in solution. *Physical Review E* **96**, 062406 (2017).
40. R. E. Dickerson *et al.*, The Anatomy of A-, B-, and Z-DNA. *Science* **216**, 475-485 (1982).
41. A. Wynveen, D. J. Lee, A. A. Kornyshev, S. Leikin, Helical coherence of DNA in crystals and solution. *Nucleic Acids Research* **36**, 5540–5551 (2008).
42. R. P. Goodman *et al.*, Rapid Chiral Assembly of Rigid DNA Building Blocks for Molecular Nanofabrication. *Science* **310**, 1661-1665 (2005).
43. S. Fischer *et al.*, Shape and Interhelical Spacing of DNA Origami Nanostructures Studied by Small-Angle X-ray Scattering. *Nano Letters* **16**, 4282-4287 (2016).

44. E. Poppleton, R. Romero, A. Mallya, L. Rovigatti, P. Sulc, OxDNA.org: a public webserver for coarse-grained simulations of DNA and RNA nanostructures. *Nucleic Acids Research* **49**, W491-W498 (2021).
45. A. A. Istratov, O. F. Vyvenko, Exponential analysis in physical phenomena. *Review of Scientific Instruments* **70**, 1233–1257 (1999).
46. R. J. L. Townshend *et al.*, Geometric deep learning of RNA structure. *Science* **373**, 1047-1051 (2021).
47. K. Zhang *et al.*, Cryo-EM structure of a 40 kDa SAM-IV riboswitch RNA at 3.7 Å resolution. *Nature Communications* **10**, 5511 (2019).
48. A. D. Garst, A. Héroux, R. P. Rambo, R. T. Batey, Crystal Structure of the Lysine Riboswitch Regulatory mRNA Element. *Journal of Biological Chemistry* **283**, 22347-22351 (2008).
49. C. D. Stoddard *et al.*, Free State Conformational Sampling of the SAM-I Riboswitch Aptamer Domain. *Structure* **18**, 787-797 (2010).
50. W. Palau, C. Di Primo, Single-cycle kinetic analysis of ternary DNA complexes by surface plasmon resonance on a decaying surface. *Biochimie* **94**, 1891-1899 (2012).
51. I. Jarmoskaite, I. AlSadhan, P. P. Vaidyanathan, D. Herschlag, How to measure and evaluate binding affinities. *eLife* **9**, e57264 (2020).
52. J. SantaLucia, D. Hicks, The thermodynamics of DNA structural motifs. *Annu Rev Bioph Biom* **33**, 415-440 (2004).
53. H. Wilson, M. Lee, Q. Wang, Probing DNA-protein interactions using single-molecule diffusivity contrast. *Biophysical Reports* **1**, 100009 (2021).
54. C. J. Barnstable *et al.*, Production of monoclonal antibodies to group A erythrocytes, HLA and other human cell surface antigens-new tools for genetic analysis. *Cell* **14**, 9-20 (1978).
55. M. N. Hug *et al.*, HLA antibody affinity determination: From HLA-specific monoclonal antibodies to donor HLA specific antibodies (DSA) in patient serum. *HLA* **102**, 278-300 (2023).
56. M. M. Schneider *et al.*, Microfluidic antibody affinity profiling of alloantibody-HLA interactions in human serum. *Biosensors and Bioelectronics* **228**, 115196 (2023).
57. W. Yoshida *et al.*, Selection of DNA aptamers against insulin and construction of an aptameric enzyme subunit for insulin sensing. *Biosensors and Bioelectronics* **24**, 1116-1120 (2009).
58. Z. Hao *et al.*, Real-Time Monitoring of Insulin Using a Graphene Field-Effect Transistor Aptameric Nanosensor. *ACS Applied Materials & Interfaces* **9**, 27504–27511 (2017).
59. D. Wu *et al.*, Flow-Cell-Based Technology for Massively Parallel Characterization of Base-Modified DNA Aptamers. *Analytical Chemistry* **95**, 2645–2652 (2023).
60. J. G. Menting *et al.*, How insulin engages its primary binding site on the insulin receptor. *Nature* **493**, 241-245 (2013).
61. T. Gutmann *et al.*, Cryo-EM structure of the complete and ligand-saturated insulin receptor ectodomain. *Journal of Cell Biology* **219**, e201907210 (2020).
62. P. De Meyts, J. Whittaker, Structural biology of insulin and IGF1 receptors: implications for drug design. *Nature Reviews Drug Discovery* **1**, 769-783 (2002).
63. V. V. Kiselyov, S. Versteyhe, L. Gauguin, P. D. Meyts, Harmonic oscillator model of the insulin and IGF1 receptors' allosteric binding and activation. *Molecular Systems Biology* **5**, 243 (2009).
64. T. Gutmann, K. H. Kim, M. Grzybek, T. Walz, Ü. Coskun, Visualization of ligand-induced transmembrane signaling in the full-length human insulin receptor. *Journal of Cell Biology* **217**, 1643–1649 (2018).

65. I. Croll, Tristan *et al.*, Higher-Resolution Structure of the Human Insulin Receptor Ectodomain: Multi-Modal Inclusion of the Insert Domain. *Structure* **24**, 469-476 (2016).
66. X. Xiong *et al.*, Symmetric and asymmetric receptor conformation continuum induced by a new insulin. *Nature Chemical Biology* **18**, 511-519 (2022).
67. P. A. Hoyne *et al.*, High affinity insulin binding by soluble insulin receptor extracellular domain fused to a leucine zipper. *FEBS Letters* **479**, 15-18 (2000).
68. F. Weis *et al.*, The signalling conformation of the insulin receptor ectodomain. *Nature Communications* **9**, 4420 (2018).
69. F. Ruggeri, M. Krishnan, Spectrally resolved single-molecule electrometry. *The Journal of Chemical Physics* **148**, 123307 (2018).
70. F. Kilár, I. Simon, The effect of iron binding on the conformation of transferrin. A small angle x-ray scattering study. *Biophysical Journal* **48**, 799-802 (1985).
71. H. Hwang, S. Myong, Protein induced fluorescence enhancement (PIFE) for probing protein–nucleic acid interactions. *Chemical Society Reviews* **43**, 1221-1229 (2014).
72. A. Langer *et al.*, A New Spectral Shift-Based Method to Characterize Molecular Interactions. *ASSAY and Drug Development Technologies* **20**, 83-94 (2022).
73. M. T. Marty *et al.*, Bayesian Deconvolution of Mass and Ion Mobility Spectra: From Binary Interactions to Polydisperse Ensembles. *Analytical Chemistry* **87**, 4370–4376 (2015).
74. M. Krishnan, N. Mojarad, P. Kukura, V. Sandoghdar, Geometry-induced electrostatic trapping of nanometric objects in a fluid. *Nature* **467**, 692-695 (2010).
75. J. R. Taylor, *An introduction to error analysis : the study of uncertainties in physical measurements*. (University Science Books, Sausalito, Calif, ed. 2nd, 1997).
76. G. Emilsson *et al.*, Strongly Stretched Protein Resistant Poly(ethylene glycol) Brushes Prepared by Grafting-To. *ACS Applied Materials & Interfaces* **7**, 7505-7515 (2015).
77. Boussinesq, Mémoire sur l'influence des frottements dans les mouvements réguliers des fluides. *Journal de Mathématiques Pures et Appliquées* **13**, 377-424 (1868).
78. S. H. Banyard, D. K. Stammers, P. M. Harrison, Electron density map of apoferritin at 2.8-Å resolution. *Nature* **271**, 282-284 (1978).
79. T. Dertinger *et al.*, Two-Focus Fluorescence Correlation Spectroscopy: A New Tool for Accurate and Absolute Diffusion Measurements. *ChemPhysChem* **8**, 433-443 (2007).
80. S. Broersma, Viscous Force Constant for a Closed Cylinder. *The Journal of Chemical Physics* **32**, 1632–1635 (1960).
81. M. M. Tirado, J. G. De La Torre, Translational friction coefficients of rigid, symmetric top macromolecules. Application to circular cylinders. *The Journal of Chemical Physics* **71**, 2581-2587 (1979).
82. J. G. De La Torre, V. A. Bloomfield, Hydrodynamic properties of complex, rigid, biological macromolecules: theory and applications. *Quarterly Reviews of Biophysics* **14**, 81-139 (1981).
83. H. C. Berg, *Random Walks in Biology*. (Princeton University Press, 1984).
84. A. Ortega, J. García de la Torre, Hydrodynamic properties of rodlike and disklike particles in dilute solution. *The Journal of Chemical Physics* **119**, 9914–9919 (2003).
85. S. H. Koenig, Brownian motion of an ellipsoid. A correction to Perrin's results. *Biopolymers* **14**, 149-160 (1975).

86. P. V. Ruijgrok, N. R. Verhart, P. Zijlstra, A. L. Tchebotareva, M. Orrit, Brownian Fluctuations and Heating of an Optically Aligned Gold Nanorod. *Physical Review Letters* **107**, 037401 (2011).
87. T. E. Ouldridge, A. A. Louis, J. P. K. Doye, Structural, mechanical, and thermodynamic properties of a coarse-grained DNA model. *Journal of Chemical Physics* **134**, 085101 (2011).
88. E. Poppleton *et al.*, Design, optimization and analysis of large DNA and RNA nanostructures through interactive visualization, editing and molecular simulation. *Nucleic Acids Research* **48**, e72 (2020).
89. F. E. Arrighi, M. Mandel, J. Bergendahl, T. C. Hsu, Buoyant densities of DNA of mammals. *Biochemical Genetics* **4**, 367-376 (1970).
90. S. R. De Kloet, B. A. G. Andrean, Buoyant density gradient centrifugation of RNA and DNA in alkali iodide solutions. *Biochimica et Biophysica Acta (BBA) - Nucleic Acids and Protein Synthesis* **247**, 519-527 (1971).
91. H. Cramér, *Mathematical Methods of Statistics*. Goldstine Printed Materials (Princeton University Press, 1946).
92. M. Zuker, Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Research* **31**, 3406-3415 (2003).
93. J. D. Altman *et al.*, Phenotypic Analysis of Antigen-Specific T Lymphocytes. *Science* **274**, 94-96 (1996).
94. S. S. Chandran *et al.*, Immunogenicity and therapeutic targeting of a public neoantigen derived from mutated PIK3CA. *Nature Medicine* **28**, 946-957 (2022).
95. W. Yoshida *et al.*, Selection of DNA aptamers against insulin and construction of an aptameric enzyme subunit for insulin sensing. *Biosensors and Bioelectronics* **24**, 1116-1120 (2009).
96. Y. Wu, B. Midinov, R. J. White, Electrochemical Aptamer-Based Sensor for Real-Time Monitoring of Insulin. *ACS Sensors* **4**, 498-503 (2019).

References

[1] A. A. Deniz, S. Mukhopadhyay, and E. A. Lemke. Single-molecule biophysics: at the interface of biology, physics and chemistry. *Journal of The Royal Society Interface*, 5(18):15–45, 2008.

[2] P. J. Rothwell, S. Berger, O. Kensch, S. Felekyan, M. Antonik, B. M. Wöhrl, T. Restle, R. S. Goody, and C. A. M. Seidel. Multiparameter single-molecule fluorescence spectroscopy reveals heterogeneity of hiv-1 reverse transcriptase:primer/template complexes. *Proceedings of the National Academy of Sciences*, 100(4):1655–1660, 2003.

[3] H. Noji, R. Yasuda, M. Yoshida, and K. Kinosita. Direct observation of the rotation of F1-ATPase. *Nature*, 386(6622):299–302, 1997.

[4] Yuichi T., Paul J. C., Gene-Wei L., Huiyi C., Mohan B., Jeremy h., Andrew E., and X. S. Xie. Quantifying e. coli proteome and transcriptome with single-molecule sensitivity in single cells. *Science*, 329(5991):533–538, 2010.

[5] C. Lorenz, J. Forsting, R. W. Style, S. Klumpp, and S. Köster. Keratin filament mechanics and energy dissipation are determined by metal-like plasticity. *Matter*, 6(6):2019–2033, 2023.

[6] F. Kulzer, T. Xia, and M. Orrit. Single molecules as optical nanoprobes for soft and complex matter. *Angewandte Chemie International Edition*, 49(5):854–866, 2010.

[7] W. E. Moerner and L. Kador. Finding a single molecule in a haystack. *Analytical Chemistry*, 61(21):1217A–1223A, 1989.

[8] M. Orrit and J. Bernard. Single pentacene molecules detected by fluorescence excitation in a p-terphenyl crystal. *Phys. Rev. Lett.*, 65:2716–2719, 1990.

[9] E. B. Shera, N. K. Seitzinger, L. M. Davis, R. A. Keller, and S. A. Soper. Detection of single fluorescent molecules. *Chemical Physics Letters*, 174(6):553–557, 1990.

[10] C. J. Bustamante, Y. R. Chemla, S. Liu, and M. D. Wang. Optical tweezers in single-molecule biophysics. *Nature Reviews Methods Primers*, 1(1):25, 2021.

[11] R. Tapia-Rojo, M. Mora, and S. Garcia-Manyes. Single-molecule magnetic tweezers to probe the equilibrium dynamics of individual proteins at physiologically relevant forces and timescales. *Nature Protocols*, 19(6):1779–1806, 2024.

[12] B. Špačková, H. K. Moberg, J. Fritzsche, J. Tenghamn, G. Sjösten, H. Šípová Jungová, D. Albinsson, Q. Lubart, D. Van Leeuwen, F. Westerlund, D. Midtvedt, E. K. Esbjörner, M. Käll, G. Volpe, and C. Langhammer. Label-free nanofluidic scattering microscopy of size and mass of single diffusing molecules and nanoparticles. *Nature Methods*, 19(6):751–758, 2022.

[13] J. Ortega Arroyo, J. Andrecka, K. M. Spillane, N. Billington, Y. Takagi, J. R. Sellers, and P. Kukura. Label-free, all-optical detection, imaging, and tracking of a single protein. *Nano Letters*, 14(4):2065–2070, 2014.

[14] J. Clarke, H.-C. Wu, L. Jayasinghe, A. Patel, S. Reid, and H. Bayley. Continuous base identification for single-molecule nanopore DNA sequencing. *Nature Nanotechnology*, 4(4):265–270, 2009.

[15] T. Gutmann, I. B. Schäfer, C. Poojari, B. Brankatschk, I. Vattulainen, M. Strauss, and U. Coskun. Cryo-EM structure of the complete and ligand-saturated insulin receptor ectodomain. *Journal of Cell Biology*, 219(1):e201907210, 2019.

[16] M. P. Backlund, R. Joyner, K. Weis, and W. E. Moerner. Correlations of three-dimensional motion of chromosomal loci in yeast revealed by the double-helix point spread function microscope. *Molecular Biology of the Cell*, 25(22):3619–3629, 2014.

[17] A. Sarkar, A. Sharma, D. Goswami, V. Namboodiri, J. Enderlein, and M. Kumbhakar. Binding constant determined from the angstrom-scale change in hydrodynamic radius of transferrin upon binding with europium using dual-focus fluorescence correlation spectroscopy. *The Journal of Physical Chemistry Letters*, 11(3):1148–1153, 2020.

[18] T. Ha, A. Y. Ting, J. Liang, W. B. Caldwell, A. A. Deniz, D. S. Chemla, P. G. Schultz, and S. Weiss. Single-molecule fluorescence spectroscopy of enzyme conformational dynamics and cleavage mechanism. *Proceedings of the National Academy of Sciences*, 96(3):893–898, 1999.

[19] T. D. Craggs, M. Sustarsic, A. Plochowietz, M. Mosayebi, H. Kaju, A. Cuthbert, J. Hohlbein, L. Domicevica, P. C. Biggin, J. P. K. Doye, and A. N. Kapanidis. Substrate conformational dynamics facilitate structure-specific recognition of gapped DNA by DNA polymerase. *Nucleic Acids Research*, 47(20):10788–10800, 09 2019.

[20] G. Young, N. Hundt, D. Cole, A. Fineberg, J. Andrecka, A. Tyler, A. Olerinyova, A. Ansari, E. G. Marklund, M. P. Collier, S. A. Chandler, O. Tkachenko, J. Allen, M. Crispin, N. Billington, Y. Takagi, J. R. Sellers, C. Eichmann, P. Selenko, L. Frey, R. Riek, M. R. Galpin, W. B. Struwe, J. L. P. Benesch, and P. Kukura. Quantitative mass imaging of single biological macromolecules. *Science*, 360(6387):423–427, 2018.

[21] F. Shi, Q. Zhang, P. Wang, H. Sun, J. Wang, X. Rong, M. Chen, C. Ju, F. Reinhard, H. Chen, J. Wrachtrup, J. Wang, and J. Du. Single-protein spin resonance spectroscopy under ambient conditions. *Science*, 347(6226):1135–1138, 2015.

[22] E. V. Yates, T. Müller, L. Rajah, E. J. De Genst, P. Arosio, S. Linse, M. Vendruscolo, C. M. Dobson, and T. P. J. Knowles. Latent analysis of unmodified biomolecules and their complexes in solution with attomole detection sensitivity. *Nature Chemistry*, 7(10):802–809, 2015.

[23] G. Krainer, R. P. B. Jacquat, M. M. Schneider, T. J. Welsh, J. Fan, Q. A. E. Peter, E. A. Andrzejewska, G. Šneiderienė, M. A. Czekalska, H. Ausserwoeger, L. Chai, W. E. Arter, K. L. Saar, T. W. Herling, T. M. Franzmann, V. Kosmoliaptsis, S. Alberti, F. U. Hartl, S. F. Lee, and T. P. J. Knowles. Single-molecule digital sizing of proteins in solution. *Nature Communications*, 15(1):7740, 2024.

[24] D. Rodriguez-Larrea and H. Bayley. Multistep protein unfolding during nanopore translocation. *Nature Nanotechnology*, 8(4):288–295, 2013.

[25] E. C. Yusko, B. R. Bruhn, O. M. Eggenberger, J. Houghtaling, R. C. Rollings, N. C. Walsh, S. Nandivada, M. Pindrus, A. R. Hall, D. Sept, J. Li, D. S. Kalonia, and M. Mayer. Real-time shape approximation and fingerprinting of single proteins using a nanopore. *Nature Nanotechnology*, 12(4):360–367, 2017.

[26] E. Erben, B. Seelbinder, I. D. Stoev, S. Klykov, N. Maghelli, and M. Kreysing. Feedback-based positioning and diffusion suppression of particles via optical control of thermoviscous flows. *Opt. Express*, 29(19):30272–30283, 2021.

[27] M. I. Bespalova, S. Mahanta, and M. Krishnan. Single-molecule trapping and measurement in solution. *Current Opinion in Chemical Biology*, 51:113–121, 2019.

[28] M. Friedel, A. Baumketner, and J.-E. Shea. Effects of surface tethering on protein folding mechanisms. *Proceedings of the National Academy of Sciences*, 103(22):8396–8401, 2006.

[29] Q. Wang, R. H. Goldsmith, Y. Jiang, S. D. Bockenhauer, and W. E. Moerner. Probing single biomolecules in solution using the anti-brownian electrokinetic (abel) trap. *Accounts of Chemical Research*, 45(11):1955–1964, 2012.

[30] J. F. Lesoine, P. A. Venkataraman, P. C. Maloney, M. E. Dumont, and L. Novotny. Nanochannel-based single molecule recycling. *Nano Letters*, 12(6):3273–3278, 2012.

[31] S. Schmid, P. Stömmer, H. Dietz, and C. Dekker. Nanopore electro-osmotic trap for the label-free study of single proteins and their conformations. *Nature Nanotechnology*, 16(11):1244–1250, 2021.

[32] P. Eberle, C. Höller, P. Müller, M. Suomalainen, U. F. Greber, H. Eghlidi, and D. Poulikakos. Single entity resolution valving of nanoscopic species in liquids. *Nature Nanotechnology*, 13(7):578–582, 2018.

[33] H. Wilson, M. Lee, and Q. Wang. Probing dna-protein interactions using single-molecule diffusivity contrast. *Biophysical Reports*, 1(1):100009, 2021.

[34] H. Wilson and Q. Wang. ABEL-FRET: tether-free single-molecule FRET with hydrodynamic profiling. *Nature Methods*, 18(7):816–820, 2021.

[35] J. A. L. Howard, B. Ambrose, M. A. S. Abdelhamid, A. Alevropoulos-Borrill, A. Ejaz, L. Dresser, M. Dienerowitz, S. D. Quinn, A. H. Squires, A. Noy, T. D. Craggs, and M. C. Leake. Rep structures can be tuned by ionicity via metastable intermediates in the absence of dna. *bioRxiv*, 2024.

[36] D. Yoo, K. L. Gurunatha, H.-K. Choi, D. A. Mohr, C. T. Ertsgaard, R. Gordon, and S.-H. Oh. Low-power optical trapping of nanoparticles and proteins with resonant coaxial nanoaperture using 10 nm gap. *Nano Letters*, 18(6):3637–3642, 2018.

[37] M. Braun, A. P. Bregulla, K. Günther, M. Mertig, and F. Cichos. Single molecules trapped by dynamic inhomogeneous temperature fields. *Nano Letters*, 15(8):5499–5505, 2015.

[38] Y. Pang and R. Gordon. Optical trapping of a single protein. *Nano Letters*, 12(1):402–406, 2012.

[39] F. Ruggeri, F. Zosel, N. Mutter, M. Różycka, M. Wojtas, A. Ożyhar, B. Schuler, and M. Krishnan. Single-molecule electrometry. *Nature Nanotechnology*, 12:488–495, 2017.

[40] R. P. B. Jacquat, G. Krainer, Q. A. E. Peter, A. N. Babar, O. Vanderpoorten, C. K. Xu, T. J. Welsh, C. F. Kaminski, U. F. Keyser, J. J. Baumberg, and T. P. J. Knowles. Single-molecule sizing through nanocavity confinement. *Nano Letters*, 23(5):1629–1636, 2023.

[41] S. Tyagi, V. VanDelinder, N. Banterle, G. Fuertes, S. Milles, M. Agez, and E. A Lemke. Continuous throughput and long-term observation of single-molecule FRET without immobilization. *Nature Methods*, 11(3):297–300, 2014.

[42] M. J. Shon and A. E. Cohen. Mass action at the single-molecule level. *Journal of the American Chemical Society*, 134(35):14618–14623, 2012.

[43] I. I. Hosseini, S. V. Hamidi, X. Capaldi, Z. Liu, M. A. Silva Pessoa, S. Mahshid, and W. Reisner. Tunable nanofluidic device for digital nucleic acid analysis. *Nanoscale*, 16:9583–9592, 2024.

[44] M. W. Elting, S. R. Leslie, L. S. Churchman, J. Korlach, C. M. J. McFaul, J. S. Leith, M. J. Levene, A. E. Cohen, and J. A. Spudich. Single-molecule fluorescence imaging of processive myosin with enhanced background suppression using linear zero-mode waveguides (zmws) and convex lens induced confinement (clic). *Opt. Express*, 21(1):1189–1202, 2013.

[45] M. Krishnan. A simple model for electrical charge in globular macromolecules and linear polyelectrolytes in solution. *The Journal of Chemical Physics*, 146(20):205101, 2017.

[46] B. W. Ninham and V. A. Parsegian. Electrostatic potential between surfaces bearing ionizable groups in ionic equilibrium with physiologic saline solution. *Journal of Theoretical Biology*, 31(3):405–428, 1971.

[47] G. S. Manning. Limiting Laws and Counterion Condensation in Polyelectrolyte Solutions I. Colligative Properties. *The Journal of Chemical Physics*, 51(3):924–933, 1969.

[48] R. Kjellander and D. J. Mitchell. An exact but linear and poisson—boltzmann-like theory for electrolytes and colloid dispersions in the primitive model. *Chemical Physics Letters*, 200(1):76–82, 1992.

[49] L. Belloni. Ionic condensation and charge renormalization in colloidal suspensions. *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, 140(1):227–243, 1998.

[50] R. R. Netz and H. Orland. Variational charge renormalization in charged systems. *The European Physical Journal E*, 11:301–311, 2003.

[51] M. Aubouy, E. Trizac, and L. Bocquet. Effective charge versus bare charge: an analytical estimate for colloids in the infinite dilution limit. *Journal of Physics A: Mathematical and General*, 36(22):5835, 2003.

[52] G. S. Manning. Electrostatic free energies of spheres, cylinders, and planes in counterion condensation theory with some applications. *Macromolecules*, 40(22):8071–8081, 2007.

[53] R. Walker-Gibbons, X. Zhu, A. Behjatian, T. J. D. Bennett, and M. Krishnan. Sensing the structural and conformational properties of single-stranded nucleic acids using electrometry and molecular simulations. *Scientific Reports*, 14:20582, 2024.

[54] M. Bespalova, A. Behjatian, N. Karedla, R. Walker-Gibbons, and M. Krishnan. Opto-electrostatic determination of nucleic acid double-helix dimensions and the structure of the molecule–solvent interface. *Macromolecules*, 55(14):6200–6210, 2022.

[55] R. J. L. Townshend, S. Eismann, A. M. Watkins, R. Rangan, M. Karelina, R. Das, and R. O. Dror. Geometric deep learning of rna structure. *Science*, 373(6558):1047–1051, 2021.

[56] J. Ma, J. Gao, J. Wang, and A. Xie. Prion-like mechanisms in parkinson's disease. *Frontiers in Neuroscience*, 13, 2019.

[57] F. Ruggeri and M. Krishnan. Spectrally resolved single-molecule electrometry. *The Journal of Chemical Physics*, 148(12):123307, 2017.

[58] H.-X. Zhou and X. Pang. Electrostatic interactions in protein structure, folding, binding, and condensation. *Chemical Reviews*, 118(4):1691–1741, 2018.

[59] Z. Li, K. C. Chan, J. D. Nickels, and X. Cheng. Electrostatic contributions to the binding free energy of nicotine to the acetylcholine binding protein. *The Journal of Physical Chemistry B*, 126(43):8669–8679, 2022.

[60] S. Ekesan, E. McCarthy, D. A. Case, and D. M. York. Rna electrostatics: How ribozymes engineer active sites to enable catalysis. *The Journal of Physical Chemistry B*, 126(32):5982–5990, 2022.

[61] F. J. Gisdon, E. Bombarda, and G. M. Ullmann. Serine and cysteine peptidases: So similar, yet different. how the active-site electrostatics facilitates different reaction mechanisms. *The Journal of Physical Chemistry B*, 126(22):4035–4048, 2022.

[62] V. V. Loladze, B. Ibarra-Molero, J. M. Sanchez-Ruiz, and G. I. Makhatadze. Engineering a thermostable protein via optimization of charge-charge interactions on the protein surface. *Biochemistry*, 38(50):16419–16423, 1999.

[63] J. M. Sanchez-Ruiz and G. I. Makhatadze. To charge or not to charge? *Trends in Biotechnology*, 19(4):132–135, 2001.

[64] A. Bah, R. M. Vernon, Z. Siddiqui, M. Krzeminski, R. Muhandiram, C. Zhao, N. Sonenberg, L. E. Kay, and J. D. Forman-Kay. Folding of an intrinsically disordered protein by phosphorylation as a regulatory switch. *Nature*, 519(7541):106–109, 2015.

[65] S. Mirceta, A. V. Signore, J. M. Burns, A. R. Cossins, K. L. Campbell, and M. Berenbrink. Evolution of mammalian diving capacity traced by myoglobin net surface charge. *Science*, 340(6138):1234192, 2013.

[66] K. M. Lebold and R. B. Best. Tuning formation of protein–dna coacervates by sequence and environment. *The Journal of Physical Chemistry B*, 126(12):2407–2419, 2022.

[67] M. V. Aquino Queirós and W. Loh. How to predict the order of phase separation of polyelectrolyte complexes and their miscibility. *The Journal of Physical Chemistry B*, 126(28):5362–5373, 2022.

[68] M. Krishnan, N. Mojarad, P. Kukura, and V. Sandoghdar. Geometry-induced electrostatic trapping of nanometric objects in a fluid. *Nature*, 467:692–695, 2010.

[69] R. Blossey. *The Poisson-Boltzmann Equation. An Introduction.* Springer Cham, 2023.

[70] H.A. Kramers. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica*, 7(4):284–304, 1940.

[71] F. Ruggeri and M. Krishnan. Lattice diffusion of a single molecule in solution. *Phys. Rev. E*, 96:062406, 2017.

[72] J. Stetefeld, S. A. McKenna, and T. R. Patel. Dynamic light scattering: a practical guide and applications in biomedical sciences. *Biophysical Reviews*, 8(4):409–427, 2016.

[73] T. Dertinger, V. Pacheco, I. von der Hocht, R. Hartmann, I. Gregor, and J. Enderlein. Two-focus fluorescence correlation spectroscopy: A new tool for accurate and absolute diffusion measurements. *ChemPhysChem*, 8(3):433–443, 2007.

[74] G. Kloes, T. J. D. Bennett, A. Chapet-Batlle, A. Behjatian, A. J. Turberfield, and M. Krishnan. Far-field electrostatic signatures of macromolecular 3d conformation. *Nano Letters*, 22(19):7834–7840, 2022.

[75] F. Ruggeri and M. Krishnan. Entropic trapping of a singly charged molecule in solution. *Nano Letters*, 18(6):3773–3779, 2018.

[76] N. Gamage, A. Barnett, N. Hempel, R. G. Duggleby, K. F. Windmill, J. L. Martin, and M. E. McManus. Human Sulfotransferases and Their Role in Chemical Metabolism. *Toxicological Sciences*, 90(1):5–22, 2005.

[77] D. M Townsend. S-glutathionylation: indicator of cell stress and regulator of the unfolded protein response. *Molecular interventions*, 7(6):313–324, 2007.

[78] K S Kosik, C L Joachim, and D J Selkoe. Microtubule-associated protein tau (tau) is a major antigenic component of paired helical filaments in alzheimer disease. *Proceedings of the National Academy of Sciences of the United States of America*, 83(11):4044–4048, 1986.

[79] J. X. Meng, Y. Zhang, D. Saman, A. M. Haider, S. De, J. C. Sang, K. Brown, K. Jiang, J. Humphrey, L. Julian, E. Hidari, S. F. Lee, G. Balmus, R. A. Floto, C. E. Bryant, J. L. P. Benesch, Y. Ye, and D. Klenerman. Hyperphosphorylated tau self-assembles into amorphous aggregates eliciting tlr4-dependent responses. *Nature Communications*, 13(1):2692, 2022.

[80] S. Verma and N. N. Nair. A comprehensive study of factors affecting the prediction of the pka shift of asp26 in thioredoxin protein. *The Journal of Physical Chemistry B*, 128(30):7304–7312, 2024.

[81] T. Hunter. Why nature chose phosphate to modify proteins. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1602):2513–2516, 2012.

[82] H. J. Reich. Introducing jce cheminfo: Organic. *Journal of Chemical Education*, 82(3):495, 2005.

[83] G. Bulaj, T. Kortemme, and D. P. Goldenberg. Ionization-reactivity relationships for cysteine thiols in polypeptides. *Biochemistry*, 37(25):8965–8972, 1998.

[84] J. P. Guthrie. Hydrolysis of esters of oxy acids: pka values for strong acids; brønsted relationship for attack of water at methyl; free energies of hydrolysis of esters of oxy acids; and a linear relationship between free energy of hydrolysis and pka holding over a range of 20 pk units. *Canadian Journal of Chemistry*, 56(17):2342–2354, 1978.

[85] A. M. Darlington and J. M. Gibbs-Davis. Bimodal or trimodal? the influence of starting ph on site identity and distribution at the low salt aqueous/silica interface. *The Journal of Physical Chemistry C*, 119(29):16560–16567, 2015.

[86] S. Wang, R. Walker-Gibbons, B. Watkins, M. Flynn, and M. Krishnan. A charge-dependent long-ranged force drives tailored assembly of matter in solution. *Nature Nanotechnology*, 19(4):485–493, 2024.

[87] C. Macias-Romero, I. Nahalka, H. I. Okur, and S. Roke. Optical imaging of surface chemistry and dynamics in confinement. *Science*, 357(6353):784–788, 2017.

[88] K. Leung, I. M. B. Nielsen, and L. J. Criscenti. Elucidating the bimodal acid-base behavior of the water-silica interface from first principles. *Journal of the American Chemical Society*, 131(51):18358–18365, 2009.

[89] M. Sulpizi, M.-P. Gaigeot, and M. Sprik. The silica-water interface: How the silanols determine the surface acidity and modulate the water properties. *Journal of Chemical Theory and Computation*, 8(3):1037–1047, 2012.

[90] A. Cimas, F. Tielens, M. Sulpizi, M.-P. Gaigeot, and D. Costa. The amorphous silica–liquid water interface studied by ab initio molecular dynamics (aimd): local organization in global disorder. *Journal of Physics: Condensed Matter*, 26(24):244106, 2014.

[91] A. Rimola, D. Costa, M. Sodupe, J.-F. Lambert, and P. Ugliengo. Silica surface features and their role in the adsorption of biomolecules: Computational modeling and experiments. *Chemical Reviews*, 113(6):4216–4313, 2013.

[92] Z. Cheng, H. Shan, Y. Sun, L. Zhang, H. Jiang, and C. Li. Evolution mechanism of surface hydroxyl groups of silica during heat treatment. *Applied Surface Science*, 513:145766, 2020.

[93] F. Ruggeri. *Escape-time electrometry: high precision measurement of electrical charge on single molecules in solution.* University of Zurich, 2018.

[94] S. C. Stein and J. Thiart. TrackNTrace: A simple and extendable open-source framework for developing single-molecule localization and tracking algorithms. *Scientific Reports*, 6(1):37947, 2016.

[95] Jean-Yves T., Nick P., Johannes S., Genevieve M. H., Gregory D. R., Emmanuel L., Sebastian Y. B., Spencer L. S., and Kevin W. E. Trackmate: An open and extensible platform for single-particle tracking. *Methods*, 115:80–90, 2017.

[96] Q. Zheng, M. F. Juette, S. Jockusch, M. R. Wasserman, Z. Zhou, R. B. Altman, and S. C. Blanchard. Ultra-stable organic fluorophores for single-molecule research. *Chem. Soc. Rev.*, 43:1044–1056, 2014.

[97] R. Roy, S. Hohng, and T. Ha. A practical guide to single-molecule FRET. *Nature Methods*, 5(6):507–516, 2008.

[98] J. Vogelsang, R. Kasper, C. Steinhauer, B. Person, M. Heilemann, M. Sauer, and P. Tinnefeld. A reducing and oxidizing system minimizes photobleaching and blinking of fluorescent dyes. *Angewandte Chemie International Edition*, 47(29):5465–5469, 2008.

[99] R. B. Altman, D. S. Terry, Z. Zhou, Q. Zheng, P. Geggier, R. A. Kolster, Y. Zhao, J. A. Javitch, J. D. Warren, and S. C. Blanchard. Cyanine fluorophore derivatives with enhanced photostability. *Nature Methods*, 9(1):68–71, 2012.

[100] R. B. Schoch, J. Han, and P. Renaud. Transport phenomena in nanofluidics. *Rev. Mod. Phys.*, 80:839–883, 2008.

[101] A. Frutiger, A. Tanno, S. Hwu, R. F. Tiefenauer, J. Vörös, and N. Nakatsuka. Nonspecific binding—fundamental concepts and consequences for biosensing applications. *Chemical Reviews*, 121(13):8095–8160, 2021.

[102] Pouran Pourhakkak, Ali Taghizadeh, Mohsen Taghizadeh, Mehrorang Ghaedi, and Sepahdar Haghdoust. Chapter 1 - fundamentals of adsorption technology. In Mehrorang Ghaedi, editor, *Adsorption: Fundamental Processes and Applications*, volume 33 of *Interface Science and Technology*, pages 1–70. Elsevier, 2021.

[103] M. Murphy and N. Dankbar. Comparison of biomolecular interaction techniques — SPR vs ITC vs MST vs BLI. https://www.reichertspr.com/insights/white-papers/comparison-of-biomolecular-interaction-techniques---white-paper. Accessed: 03-10-2024.

[104] M. Götz, A. Barth, D. S.-R. Bohr, R. Börner, J. Chen, T. Cordes, D. A. Erie, C. Gebhardt, M. C. A. S. Hadzic, G. L. Hamilton, N. S. Hatzakis, T. Hugel, L. Kisley, D. C. Lamb, C. De Lannoy, C. Mahn, D. Dunukara, D. De Ridder, H. Sanabria, J. Schimpf, C. A. M. Seidel, R. K. O. Sigel, M. B. Sletfjerding, J. Thomsen, L. Vollmar, S. Wanninger, K. R. Weninger, P. Xu, and S. Schmid. A blind benchmark of analysis tools to infer kinetic rate constants from single-molecule FRET trajectories. *Nature Communications*, 13(1):5402, 2022.

[105] E. Lerner, A. Barth, J. Hendrix, B. Ambrose, V. Birkedal, S. C. Blanchard, R. Börner, H. Sung Chung, T. Cordes, T. D. Craggs, A. A. Deniz, J. Diao, J. Fei, R. L. Gonzalez, I. V. Gopich, T. Ha, C. A. Hanke, G. Haran, N. S. Hatzakis, S. Hohng, S.-C. Hong, T. Hugel, A. Ingargiola, C. Joo, A. N. Kapanidis, H. D. Kim, T. Laurence, N. K. Lee, T.-H. Lee, E. A. Lemke, E. Margeat, J. Michaelis, X. Michalet, S. Myong, D. Nettels, T.-O. Peulen, E. Ploetz, Y. Razvag, N. C. Robb, B. Schuler, H. Soleimaninejad, C. Tang, R. Vafabakhsh, D. C. Lamb, C. A. M. Seidel, and S. Weiss. Fret-based dynamic structural biology: Challenges, perspectives and an appeal for open-science practices. *eLife*, 10:e60416, mar 2021.

[106] T. Gilboa, P. M. Garden, and L. Cohen. Single-molecule analysis of nucleic acid biomarkers – a review. *Analytica Chimica Acta*, 1115:61–85, 2020.

[107] D. M. Rissin, C. W. Kan, T. G. Campbell, S. C. Howes, D. R. Fournier, L. Song, T. Piech, P. P. Patel, L. Chang, A. J. Rivnak, E. P. Ferrell, J. D. Randall, G. K. Provuncher, D. R. Walt, and D. C. Duffy. Single-molecule enzyme-linked immunosorbent assay detects serum proteins at subfemtomolar concentrations. *Nature Biotechnology*, 28(6):595–599, 2010.

[108] C. Reily, T. J. Stewart, M. B. Renfrow, and J. Novak. Glycosylation in health and disease. *Nature Reviews Nephrology*, 15(6):346–366, June 2019.

[109] E. O. J. Porta and P. G. Steel. Activity-based protein profiling: A graphical review. *Current Research in Pharmacology and Drug Discovery*, 5:100164, 2023.

[110] H. He, C. Wu, M. Saqib, and R. Hao. Single-molecule fluorescence methods for protein biomarker analysis. *Analytical and Bioanalytical Chemistry*, 415(18):3655–3669, 2023.

[111] Z. Williams, I. Z. Ben-Dov, R. Elias, A. Mihailovic, M. Brown, Z. Rosenwaks, and T. Tuschl. Comprehensive profiling of circulating microrna via small rna sequencing of cdna libraries reveals biomarker potential and limitations. *Proceedings of the National Academy of Sciences*, 110(11):4255–4260, 2013.

[112] C. Campos-Escamilla, D. Siliqi, L. A. Gonzalez-Ramirez, C. Lopez-Sanchez, J. A. Gavira, and A. Moreno. X-ray characterization of conformational changes of human apo- and holo-transferrin. *International Journal of Molecular Sciences*, 22(24), 2021.

[113] P. K. Choudhary and R. K. O. Sigel. Mg^{2+}-induced conformational changes in the btub riboswitch from e. coli. *RNA*, 20(1):36–45, 2014.

[114] G. Holdgate, K. Embrey, A. Milbradt, and G. Davies. Biophysical methods in early drug discovery. *ADMET and DMPK*, 7(4):222–241, 2019.

[115] S. Mahanta, P. Vallejo-Ramirez, N. Karedla, P. Puczkarski, and M. Krishnan. Wide-field optical imaging of electrical charge and chemical reactions at the solid–liquid interface. *Proceedings of the National Academy of Sciences*, 119(49):e2209955119, 2022.