

# Real-time shape approximation and fingerprinting of single proteins using a nanopore

Erik C. Yusko<sup>1†</sup>, Brandon R. Bruhn<sup>1†</sup>, Olivia M. Eggenberger<sup>1,2</sup>, Jared Houghtaling<sup>1,2</sup>, Ryan C. Rollings<sup>3</sup>, Nathan C. Walsh<sup>3</sup>, Santoshi Nandivada<sup>3</sup>, Mariya Pindrus<sup>4</sup>, Adam R. Hall<sup>5</sup>, David Sept<sup>1,6</sup>, Jiali Li<sup>3</sup>, Devendra S. Kalonia<sup>4</sup> and Michael Mayer<sup>1,2,7\*</sup>

**Established methods for characterizing proteins typically require physical or chemical modification steps or cannot be used to examine individual molecules in solution. Ionic current measurements through electrolyte-filled nanopores can characterize single native proteins in an aqueous environment, but currently offer only limited capabilities. Here we show that the zeptolitre sensing volume of bilayer-coated solid-state nanopores can be used to determine the approximate shape, volume, charge, rotational diffusion coefficient and dipole moment of individual proteins. To do this, we developed a theory for the quantitative understanding of modulations in ionic current that arise from the rotational dynamics of single proteins as they move through the electric field inside the nanopore. The approach allows us to measure the five parameters simultaneously, and we show that they can be used to identify, characterize and quantify proteins and protein complexes with potential implications for structural biology, proteomics, biomarker detection and routine protein analysis.**

Methods to characterize and quantify unlabelled, folded proteins in aqueous environments and on a single-molecule level do not currently exist<sup>1</sup>. If available, such methods could enhance routine protein analysis, enabling rapid and sensitive biomarker detection<sup>2</sup>, and allowing the analysis of personal proteomes<sup>3</sup>. Furthermore, if these methods could provide low-resolution approximations of shape, volume and dipole moment, they could help to reveal the conformation of transient protein complexes or large assemblies that are not accessible by electron microscopy, NMR spectroscopy, X-ray crystallography or small-angle X-ray scattering<sup>4</sup>.

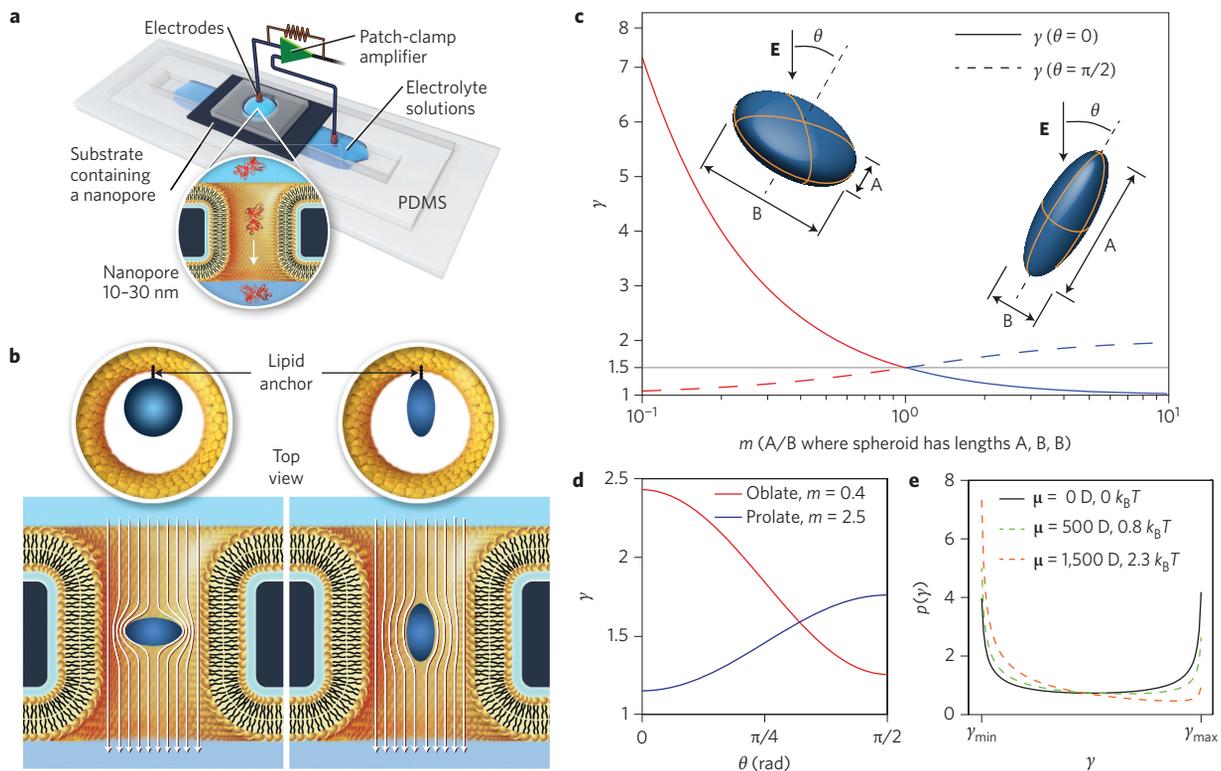
Despite the pioneering work by Oncley<sup>5</sup>, dipole moment has mostly been neglected as a protein descriptor and existing methods for determining protein dipole moments are tedious and limited to ensemble measurements. The dipole moment could, however, provide a powerful dimension for label-free protein analysis since absolute values range from zero to several thousand Debye among different proteins and are not correlated with protein size or charge<sup>6</sup>. Furthermore, the pharmaceutical industry is increasingly recognizing the importance of dipole moment for antibody formulations<sup>7</sup>, in part because subcutaneous injection of highly concentrated solutions of monoclonal antibodies (the fastest growing class of therapeutics) can be impractical due to high viscosity and aggregation resulting from dipole alignment<sup>7–9</sup>.

Nanopores can be used to interrogate single proteins. The approach requires a single electrolyte-filled pore in a thin insulating membrane that connects two solutions and can serve as a conduit for ions and proteins (Fig. 1a)<sup>10,11</sup>. Electrodes connect the solutions on both sides of the membrane to a high-gain amplifier that applies a constant electric potential difference, while measuring the ionic current through the nanopore. This arrangement ensures that the

applied voltage drops almost entirely within the pore, rendering this zone sensitive to transient changes in its ionic conductivity. Consequently, each protein that is driven electrophoretically through the pore displaces conductive electrolyte, distorts the electric field and reduces the ionic current through the pore<sup>12,13</sup>. If the volume of the electrolyte-filled pore is sufficiently small compared with the volume of the particle, then the change in ionic current due to the translocating particle is measurable and characterized by its magnitude,  $\Delta I$ , and duration,  $t_d$  (refs 12,14 15–17); this current signature is referred to as a resistive pulse. In addition to its sensitivity to conductivity changes, this small volume transiently separates single proteins from other macromolecules in solution. As we report here, this allows the rotational dynamics of individual proteins to be interrogated and interpreted based on time- and orientation-dependent modulations of ionic current (Fig. 1b–e, Supplementary Notes 1–6 and Supplementary Figs. 1–9).

Several groups have recently considered, in qualitative terms, the effect of a protein's<sup>14,16,18–21</sup> or nanoparticle's<sup>12,22</sup> shape when analysing distributed  $\Delta I$  signals<sup>23</sup> and also the effect of a protein's dipole on its translocation through an  $\alpha$ -hemolysin pore in the presence of an a.c. field<sup>24</sup>. We have now developed a quantitative understanding of the dependence of measured  $\Delta I$  values on the volume, shape, dipole moment and rotational diffusion coefficient of a protein inside a cylindrical nanopore, which makes it possible to estimate these parameters from resistive pulses (Supplementary Fig. 10). With further improvements, the ability to analyse individual proteins should mean the approximate shape of the protein, or the other four parameters, can be determined in mixtures of proteins without purification; existing methods for determining the shape or structure of proteins either require purified, concentrated, or crystallized protein samples or the protein dynamics cannot be examined.

<sup>1</sup>Department of Biomedical Engineering, University of Michigan, Ann Arbor, Michigan 48109, USA. <sup>2</sup>Adolphe Merkle Institute, University of Fribourg, Chemin des Verdiers 4, CH-1700 Fribourg, Switzerland. <sup>3</sup>Department of Physics, University of Arkansas, Fayetteville, Arkansas 72701, USA. <sup>4</sup>Department of Pharmaceutical Sciences, University of Connecticut, Storrs, Connecticut 06269, USA. <sup>5</sup>Department of Biomedical Engineering and Comprehensive Cancer Center, Wake Forest University School of Medicine, Winston Salem, North Carolina 27157, USA. <sup>6</sup>Center for Computational Medicine and Biology, University of Michigan, Ann Arbor, Michigan 48109, USA. <sup>7</sup>Biophysics Program, University of Michigan, Ann Arbor, Michigan 48109, USA. <sup>†</sup>These authors contributed equally to this work. \*e-mail: michael.mayer@unifr.ch



**Figure 1 | Rotational dynamics of individual proteins inside a nanopore reveal a spheroidal approximation of the protein's shape  $m$ .** **a**, Set-up to measure resistive pulses from the translocation of individual proteins. PDMS, polydimethylsiloxane. **b**, Top and side views of a nanopore illustrating the two extreme orientations of a spheroidal protein that is anchored to a fluid lipid coating on the pore wall. A crosswise orientation disturbs the field lines inside the pore more than a lengthwise orientation due to the angle-dependent electrical shape factor  $\gamma$  (ref. 27). All nanopores used in this work were longer than the size of the translocating proteins and the nanopore geometry was approximated as cylindrical. **c**, Electrical shape factor  $\gamma$  of spheroids (prolates in blue curves and oblates in red curves) as a function of their aspect ratio,  $m$ , for two extreme orientations: when the angle,  $\theta$ , between the axis of rotation of the ellipsoid relative to the electric field  $\mathbf{E}$  is 0, that is,  $\theta = 0$  (solid curves), and when  $\theta = \pi/2$  (dashed curves). For reference, a sphere has a  $m$  value equal to 1, and an electrical shape factor  $\gamma$  of 1.5 that is independent of its angle  $\theta$  (grey line)<sup>13,26–28</sup>. **d**, Electrical shape factor  $\gamma$  as a function of  $\theta$  for prolates with a defined  $m$  value of 2.5 and oblates with a  $m$  value of 0.4. **e**, Bimodal probability distribution of electrical shape factors,  $p(\gamma)$ , for spheroids without a dipole moment  $\mu$  as predicted by Golibersuch (black curve)<sup>13</sup> and for spheroidal proteins with a dipole moment of 500 and 1,500 Debye pointed parallel to the longest axis of the protein (dashed curves). For the different magnitudes of the dipole moment, the energy difference between  $\theta = 0$  and  $\theta = \pi/2$  is listed in units of  $k_B T$  (where  $k_B$  is Boltzmann's constant and  $T$  is temperature) for a typical electric field of  $2 \times 10^6 \text{ V m}^{-1}$ . See Supplementary Notes 2 and 9 for details.

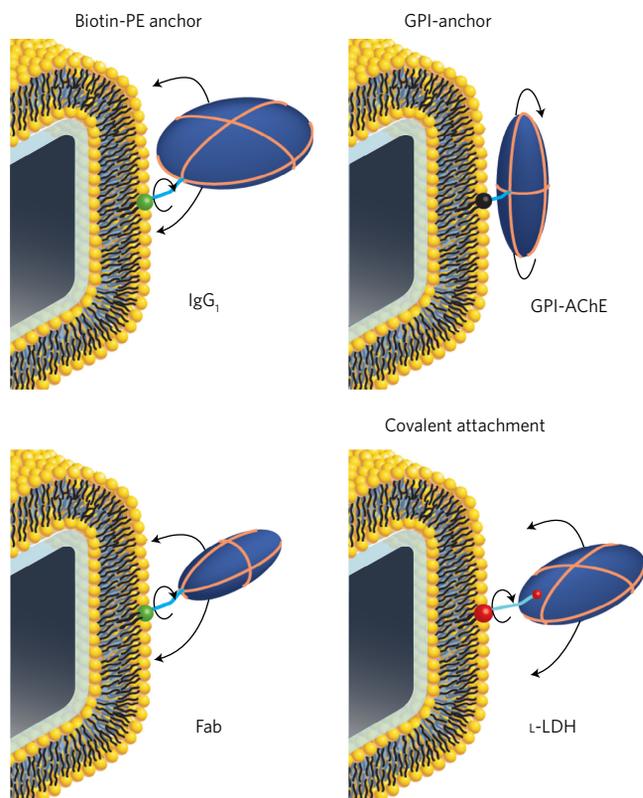
### Theory of spheroids rotating in an electric field

The main concept underlying the analysis introduced in this work is that rotation of a single non-spherical object during translocation through a cylindrical nanopore<sup>25</sup> modulates the current reduction through the pore and that these modulations can be used to determine the orientation, approximate shape and volume of the object in the pore (Fig. 1b–e). In the context of this work, we approximate the shape of proteins  $m$  with the axis ratio of a spheroid:  $m = A/B$  (Fig. 1c).

**Particle orientation.** Golibersuch<sup>13</sup> and others<sup>26,27</sup> demonstrated theoretically on ideal spheroids and experimentally on red blood cells that a crosswise orientation of an oblate or prolate disturbs the electric field along a tube more dramatically than a lengthwise orientation (Fig. 1b). In the context of current recordings through a nanopore this means that the particle-induced blockade of current,  $\Delta I$ , is maximal when the spheroidal particle is in its extreme crosswise orientation and minimal in the extreme lengthwise orientation. As shown in Fig. 1c–e, Fricke<sup>28</sup> and later Velick and Gorin<sup>29</sup> as well as Golibersuch<sup>13</sup> quantified these effects with an electrical shape factor  $\gamma$ , which is determined by the particle shape  $m$  and directly proportional to the current reduction  $\Delta I$  (Supplementary Note 2).

The black curve in Fig. 1e shows that for randomly rotating spheroids, the two extreme electrical shape factors  $\gamma_{\min}$  and  $\gamma_{\max}$  are most probable because  $\gamma$  is less angle-dependent near the extreme lengthwise and crosswise orientations than in intermediate orientations (see Supplementary Note 2 for details). This U-shaped probability distribution of  $\gamma$  means that the translocation of randomly rotating spheroidal proteins through a uniform electrical field in a cylindrical nanopore should result in a distribution of  $\Delta I$  values with two maxima, one corresponding to  $\Delta I(\gamma_{\min})$  and one to  $\Delta I(\gamma_{\max})$ . In contrast, spherical proteins, with a  $\gamma$  value of 1.5 that is independent of orientation, should result in normal distributions of  $\Delta I$  values.

**Particle shape and volume.** When comparing the translocation of two oblates of equal volume but different shape, the particle deviating most from a perfect sphere (that is, the flatter oblate) disturbs the field lines more dramatically in a crosswise orientation and less dramatically in a lengthwise orientation than the rounder object in these two orientations (Fig. 1c). In contrast, when comparing two particles with the same shape but different volume, both the orientation-dependent minimal and maximal current reductions have larger magnitudes for the larger particle compared with the smaller one. In other words, particles with increasingly non-spherical shapes result in a more extreme ratio



**Figure 2 | Three different strategies of anchoring proteins to the lipid coating used in this work to slow down translocation such that rotational diffusion of the proteins could be resolved in time.** A lipid anchor with a biotin group selectively captured anti-biotin immunoglobulin  $G_1$  ( $IgG_1$ ) antibodies and Fab fragments, an intrinsic glycosylphosphatidylinositol (GPI) anchor captured acetylcholinesterase (AChE), and a bifunctional amine-reactive crosslinker provided a general strategy to attach proteins of interest covalently to ethanolamine lipids in the bilayer coating. All proteins analysed in this work were tethered with a phospholipid anchor to the bilayer by one of these three strategies. These tethers were sufficiently long ( $\geq 1.5$  nm in their extended conformation) and flexible ( $\geq 12$   $\sigma$ -bonds) and nanopore diameters were at least twice the volume-equivalent spherical diameter of the examined proteins, such that the proteins were able to rotate and sample all possible orientations. L-LDH, L-lactate dehydrogenase.

between the current blockage in their crosswise versus lengthwise orientation, while particles with increasing volumes result in larger magnitudes of all current reductions.

### Tethering proteins enables determination of their charge

To obtain time-resolved values of  $\Delta I$  from the translocation of single proteins, we slowed down translocation by tethering proteins to a lipid anchor that was embedded in the fluid lipid bilayer coating of the nanopores (Figs 1b and 2; see Supplementary Notes 4 and 7 for a detailed discussion on the effects of lipid tethering and the point of attachment)<sup>16,21,30</sup>. In this way, the speed of protein translocation was dominated by the two orders of magnitude higher viscosity of the lipid coating compared with that of the aqueous electrolyte. In addition, we maximized the possibility that the proteins could rotate and sample all orientations in the nanopore by employing long and flexible tethers (Fig. 2). Finally, the lipid coating minimized non-specific interactions between proteins and the pore wall<sup>16</sup>, thus enabling extraction of quantitative data on Brownian rotational and translational dynamics of proteins while they were in the pore<sup>31</sup>. For instance, we took advantage of the resulting translocation times to determine the net charge of all ten proteins studied and found a

strong correlation between the charge from nanopore experiments and reference values for the charge of each protein (Pearson correlation coefficient  $r = 0.95$ , see Supplementary Fig. 11).

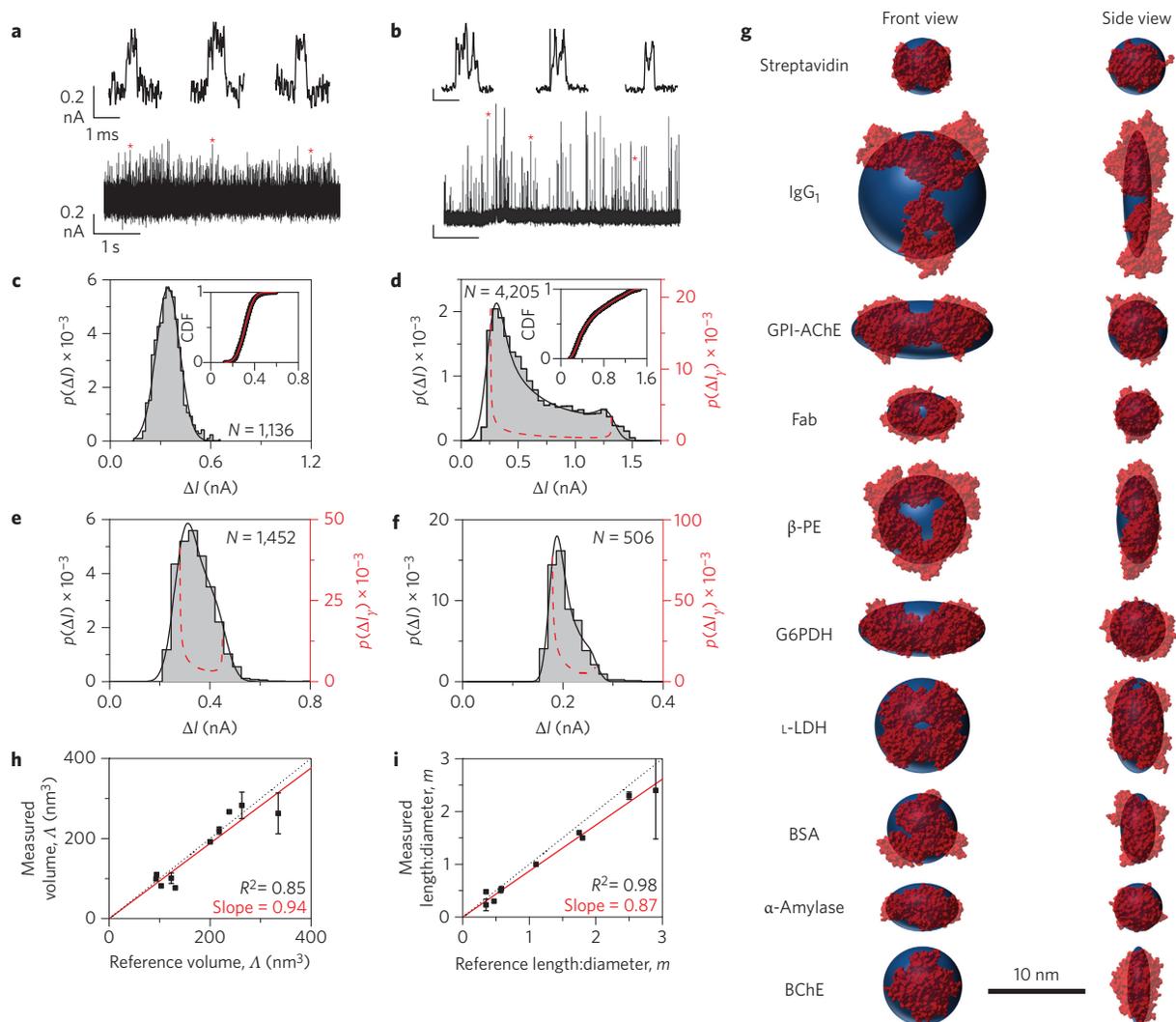
### Current blockades reveal the shape and volume of proteins

To determine the approximate shape and volume of proteins, we developed two strategies based on the theory developed by Fricke<sup>28</sup>, Velick and Gorin<sup>29</sup> and Golibersuch<sup>13</sup> (Fig. 1c–e and Supplementary Fig. 12). Both strategies approximate the shape of proteins with a spheroid in a cylindrical pore (Supplementary Fig. 13) and have different strengths and weaknesses as demonstrated in Figs 3 and 4.

The first strategy estimates shape and volume from distributions of maximum  $\Delta I$  values from many translocation events that were obtained from a pure protein sample. In other words, only the peak value of  $\Delta I$  from each resistive pulse is used for analysis. Maximum  $\Delta I$  values have been employed in almost all nanopore-based resistive pulse analyses of protein volume to date combined with the assumption of a perfectly spherical particle shape (that is,  $m = 1$  and  $\gamma = 1.5$ ), thereby forgoing the opportunity to evaluate protein shape and introducing error for particles that deviate from perfect spheres<sup>13</sup>. In contrast, we demonstrate here that Golibersuch's approach to determine the shape of cells and microparticles<sup>13</sup> can be adopted to approximate the shape of protein molecules.

Determining the shape and volume of spheroids from distributions of maximum  $\Delta I$  values proceeds in three steps; Fig. 3 shows the results from each step (see Supplementary Note 2 and Supplementary Figs 5–8 for details). First, an algorithm detects resistive pulses from the translocation of hundreds to thousands of copies of the same protein and determines the maximum amplitude of the current modulation,  $\Delta I$ , with respect to the baseline current for each pulse (Fig. 3a,b). As predicted theoretically in Fig. 1c–e, the resulting distribution of maximum  $\Delta I$  values is either normal for spherical proteins (Fig. 3c) or bimodal for non-spherical proteins (see Fig. 3d–f and Supplementary Fig. 6 and Supplementary Note 2). Second, to circumvent binning effects encountered with probability distributions<sup>16</sup>, the experimentally determined distribution of  $\Delta I$  values is converted to an empirical cumulative density function, CDF (Fig. 3c,d, insets), and fit iteratively with an equation that describes the variation in  $\Delta I$  due to rotation of proteins with non-spherical shape inside a cylindrical nanopore (Supplementary Note 2, Supplementary equation 13a,b). We refer to this equation as the convolution model since it also accounts for broadening of the  $\Delta I$  distribution due to convolution of the true signal with noise (Supplementary Fig. 5) and for bias towards either the crosswise or lengthwise orientation as a result of the electric-field-induced torque on the protein's dipole moment (Supplementary Note 9)<sup>32</sup>. The bias in a distribution of maximum  $\Delta I$  values, however, may also be affected by other factors than the dipole moment (as discussed in Supplementary Note 2), which are all accounted for by the same fitting parameter. The values of  $\Delta I(\gamma_{\min})$  and  $\Delta I(\gamma_{\max})$  returned by the fitting procedure reflect the two extreme orientations of the protein (red dashed curves in Fig. 3d–f). Third, based on the direct proportionality between  $\Delta I$  and  $\gamma$  and the geometrical relationship between  $\gamma$  and the length-to-diameter ratio  $m$  of a spheroid (Supplementary Note 2, Supplementary equations 1 and 4–7), we determine the shape and volume that agree best with the experimental distribution of  $\Delta I$  values for the protein.

Figure 3g shows the spheroidal approximation of the shape of ten different proteins compared with the respective crystal structure for each protein, illustrating that this analysis yields excellent estimates of protein shape, particularly for proteins that closely resemble a spheroid. Figure 3h,i, for instance, shows that the volume and  $m$  values agree well with the expected reference values; the average deviation of both parameters is less than 20% (Supplementary Tables 1–4 and Supplementary Fig. 14 show the results of this analysis as well as reference values). These results also show that two proteins with a similar molecular weight and volume but different

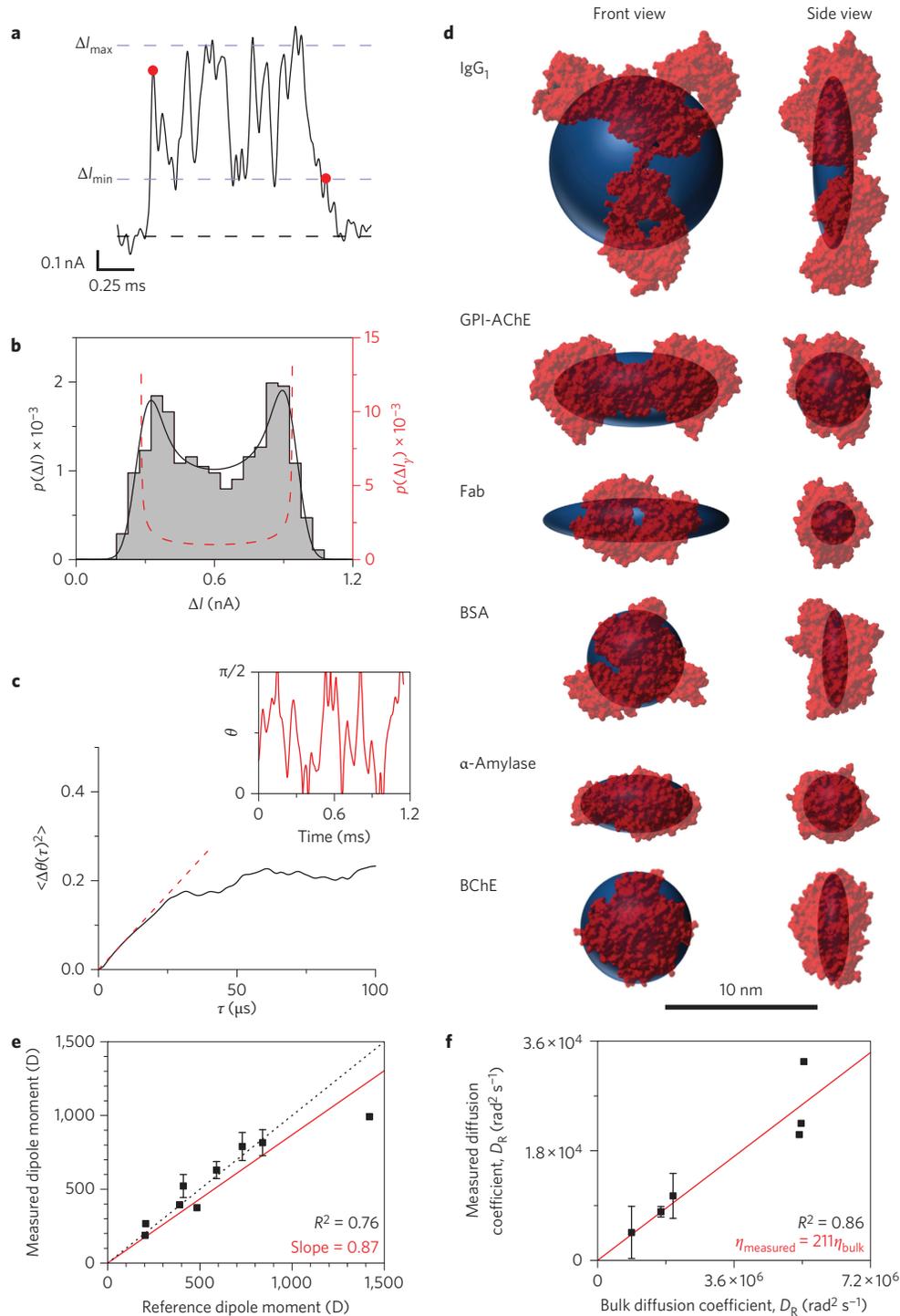


**Figure 3 | Determination of approximate protein shape and volume from histograms of maximum  $\Delta I$  values from resistive pulse recordings.** **a, b**, Examples of original current traces as a function of time: upward spikes indicate individual resistive current pulses towards zero current due to the translocation of single streptavidin (**a**) or IgG (**b**) proteins. Resistive pulses marked by an asterisk are shown in detail above. **c–f**, Histograms of maximum  $\Delta I$  values from resistive pulse recordings with streptavidin (**c**), IgG<sub>1</sub> (**d**), GPI-AChE (**e**) and G6PDH (**f**) proteins. Black curves show the solution of the convolution model,  $p(\Delta I)$ , after a nonlinear least-squares fitting procedure, and red dashed curves show the estimated distribution of  $\Delta I$  values due to the distribution of electrical shape factors,  $p(\Delta I_\gamma)$ . Supplementary Table 1 lists the values of all fitting parameters and the electric field strength used in each experiment. Supplementary Note 2 and Supplementary Figs 5–7 explain the convolution model and fitting procedure in detail and extend the analyses to all proteins characterized in this work. **g**, Comparison of the approximate shape of ten proteins as determined by analysis of resistive pulses (blue spheroids) with crystal structures from the Protein Data Bank in red (streptavidin, 3RY1; anti-biotin immunoglobulin G<sub>1</sub>, 1HZH; GPI-anchored acetylcholinesterase, 3LII; anti-biotin Fab fragment, 1F8T;  $\beta$ -phycoerythrin ( $\beta$ -PE), 3V57; glucose-6-phosphate dehydrogenase (G6PDH), 4EM5; L-lactate dehydrogenase (L-LDH), 2ZQY; bovine serum albumin (BSA), 3V03;  $\alpha$ -amylase, 1BLI; and butyrylcholinesterase (BChE), 1POI). **h**, Comparison of the measured volume by nanopore-based analysis with the expected reference volume. **i**, Comparison of the measured length-to-diameter ratios  $m$  of all proteins with the expected reference values of  $m$ . Error bars in **h, i** represent the standard deviation in most probable values from experiment-to-experiment or day-to-day.

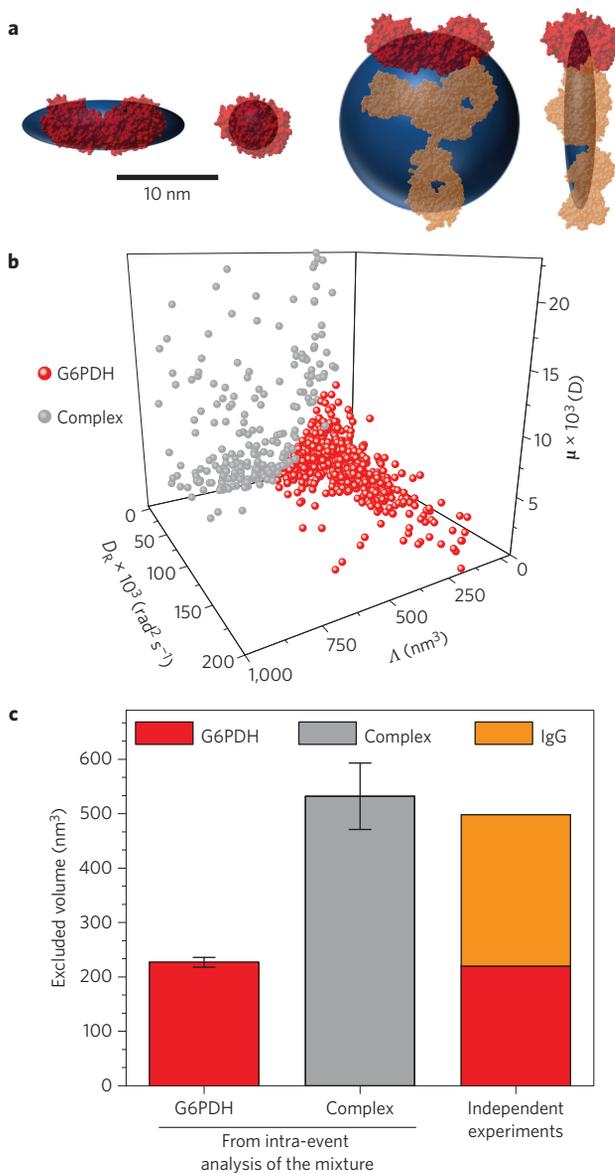
shape are clearly distinguishable by this analysis; for instance, compare the ellipsoids determined for the IgG<sub>1</sub> antibody and GPI-AChE in Fig. 3g (see also Supplementary Figs 15 and 16)<sup>23</sup>. We propose that protein shape  $m$  is an excellent protein descriptor that is not correlated with protein size or charge (Supplementary Note 8 and Supplementary Figs 17–19).

As opposed to this first strategy, which analyses maximum  $\Delta I$  values from many translocation events, the second strategy makes it possible to estimate the shape and volume of single proteins from individual resistive pulses by analysing all current values from the beginning to the end of individual translocation events,  $\Delta I(t)$  (Fig. 4a). This analysis relies on a single translocating protein to rotate and sample virtually all orientation-dependent  $\gamma$

values such that the resulting single-event, or intra-event  $\Delta I$  distribution, reveals  $\Delta I(\gamma_{\max})$  and  $\Delta I(\gamma_{\min})$  and thereby the protein's spheroidal shape approximation and its volume. Another advantage of this strategy is that it can also determine the protein's rotational diffusion coefficient and dipole moment from individual resistive pulses based on orientation-dependent modulations in current over time (Fig. 4e,f). In fact, estimates of all four parameters can be determined and updated in real time as a single protein travels through the pore (Supplementary Fig. 10). The disadvantage of this simultaneous multiparameter analysis from single molecules is that the analysis was limited to the approximately 10% of resistive pulses, which lasted at least 400  $\mu$ s to ensure sampling of the full range of electrical shape factors  $\gamma$ . We chose this duration based



**Figure 4 | Approximate shape, dipole moment and rotational diffusion coefficient obtained from current modulations within individual resistive pulses from the translocation of a single protein.** **a**, Resistive pulse from the translocation of a single IgG<sub>1</sub> molecule. Red dots mark the beginning and end of the resistive pulse as identified by an automated algorithm. **b**, Distribution of all current values within this one resistive pulse. The black curve shows the solution of the convolution model,  $p(\Delta I)$ , after a nonlinear least-squares fitting procedure, and the red dashed curve shows the estimated distribution of  $\Delta I$  values due to the distribution of electrical shape factors,  $p(\Delta I_e)$ . **c**, Curve of mean-square angular displacement (black) and initial slope of this curve (dashed red). The inset shows the time-dependent angular displacement  $\theta(t)$  of a single protein in the pore as obtained from transformation of recorded intra-event  $\Delta I(t)$  data (see Supplementary Note 6 for details). **d**, Comparison of the approximate shape of proteins as determined by analysis of individual resistive pulses (blue) with crystal structures in red (blue spheroids show the median values of  $m$  and volume from single-event analyses of each protein; see Supplementary Fig. 15 for complete distributions from the single-event analyses). **e**, The most frequently observed dipole moments (in ascending order) of G6PDH, L-lactate dehydrogenase,  $\alpha$ -amylase,  $\beta$ -phycoerythrin, bovine serum albumin, Fab, GPI-AChE, IgG<sub>1</sub> and BChE agree well with expected reference values of their dipole moments. **f**, The most frequently observed rotational diffusion coefficients (in ascending order) of IgG<sub>1</sub>,  $\beta$ -phycoerythrin, GPI-AChE, BChE, Fab and  $\alpha$ -amylase agree well with the expected reference values. The signal-to-noise ratio for G6PDH, L-lactate dehydrogenase and bovine serum albumin was too small to determine accurate values of  $D_R$ . Error bars in **e, f** represent the standard deviation in most probable values from experiment-to-experiment.



**Figure 5 | Fingerprinting of individual translocation events permits identification and characterization of G6PDH and a G6PDH-IgG complex from a binary mixture.** **a**, The volume,  $\Lambda$ , and approximate shape of G6PDH (left) and the G6PDH-IgG complex (right) as determined by analysis of individual resistive pulses is similar to the crystal structures in red. Blue spheroids show the median values of  $m$  and  $\Lambda$  determined from single-event analyses and classification of each event. **b**, Values for the volume, rotational diffusion coefficient and dipole moment determined from individual events. The kmeans clustering algorithm in MATLAB classified single events as corresponding to a single G6PDH (red points) or to the G6PDH-IgG complex (grey points) (see Supplementary Note 8 and Supplementary Fig. 22). This single-event classification estimated that 28% of events were due to the complex, which is nearly the same proportion of events estimated to be in the complex based on analysis of maximum  $\Delta I$  values from distributions of hundreds of resistive pulses (Supplementary Fig. 22). **c**, The volume of G6PDH and the G6PDH-IgG complex determined by single-event analysis and classification of events from the mixture are nearly identical (<10% deviation) to the sum of volumes obtained for G6PDH in an experiment without anti-G6PDH IgG and the volume of an individual IgG. Error bars represent the standard error of the median volume value (Supplementary Note 8).

on the mean-square angular displacement equation that predicts a tethered protein will sample all possible orientations in less than

400  $\mu$ s as long as its rotational diffusion coefficient exceeds 3,000  $\text{rad}^2 \text{s}^{-1}$  (Supplementary Table 4).

Figure 4 shows estimates of the shape and volume of proteins obtained from fitting distributions of intra-event  $\Delta I(t)$  values from individual resistive pulses with the convolution model in the same way as the distributions of maximum  $\Delta I$  values from hundreds of pulses (Supplementary Note 6 and Supplementary Fig. 20). We find that these intra-event  $\Delta I$  distributions retain their key features (for example, minimal and maximal  $\Delta I$  values) despite smoothing of the current recordings due to filtering (Supplementary Fig. 21). The median protein shapes obtained from this analysis are in reasonable agreement with their crystal structure (Fig. 4d), although the analysis of maximum  $\Delta I$  values yielded more accurate shapes (Fig. 3g,i and see Supplementary Notes 2 and 6 for discussion). With regard to the robustness of each stand-alone single-molecule measurement, more than half of all measurements yielded values of protein shape  $m$  and volume  $\Lambda$  that were within  $\pm 35\%$  of the median value (Supplementary Fig. 20), indicating that intra-event analysis has the potential to yield good estimates of shape and volume of single proteins from individual translocation events. Moreover, this strategy of analysing intra-event  $\Delta I$  distributions introduces the first method for estimating, in real time, the shape and volume of single protein molecules *in situ*. This capability is particularly advantageous for analysis of protein samples with large heterogeneity in size and shape (such as amyloids); ensemble methods such as dynamic light scattering are not well suited for such samples<sup>21</sup>. Other techniques for analysing the shape and volume of single proteins such as cryo-electron microscopy and atomic force microscopy either require freezing or surface immobilization that fixes the orientation of the proteins; therefore, these methods are not well suited for tracking protein dynamics.

### Current fluctuations reveal rotational diffusion of proteins

Figure 4 shows that monitoring the time-dependent modulations of  $\Delta I$  while a single particle moves through a nanopore makes it possible to measure its rotational diffusion coefficient,  $D_R$ , by tracking its rotation over short timescales and therefore over small fluctuations in angle (Supplementary Notes 6 and 7 and Supplementary Figs 22–24). We carried out this analysis in three steps by transforming the intra-event current signal into an angle (that is, orientation) versus time curve (Supplementary Note 6), calculating the mean-square angular displacement over various time intervals,  $\tau$ , and fitting its initial slope with a model for rotational diffusion about a single axis (Fig. 4c). Figure 4f shows that the most probable  $D_R$  values for tethered proteins obtained from many intra-event analyses of individual resistive pulses were strongly correlated with the expected values of  $D_R$  in bulk solution (Pearson's  $r = 0.93$ ). The resulting functional relationship makes it possible to estimate bulk  $D_R$  values from translocations of single proteins through nanopores in less than a millisecond. As expected, the presence of the lipid tether and close proximity of the proteins to the bilayer coating reduced  $D_R$  significantly<sup>33,34</sup>; this tether-induced attenuation of rotation was consistent with an apparent increase in viscosity  $\eta$  by a factor of 210 compared with the viscosity in bulk solution (Supplementary Fig. 22). This value is in excellent agreement with fluorescence polarization measurements of GPI-anchored AChE by Yuan and Axelrod, which revealed that the rotational diffusion coefficient of lipid-tethered AChE was 200 times smaller than its expected value in bulk solution<sup>35</sup>. Thus for analysing the rotational dynamics of single proteins in real time as introduced here, this tether-induced reduction of  $D_R$  was beneficial as it enabled changes in protein orientation to be resolved in time (Supplementary Figs 10 and 21).

With regard to the robustness of these measurements, we found that, on average, the relative standard deviation of the most probable value of  $D_R$  from distributions of measured single molecule values was 46% from experiment-to-experiment; however, as is typical for many single-molecule measurements, the variation from

event-to-event was large with a mean absolute deviation of 400% (Supplementary Fig. 22).

### Bias in protein orientation reveals its dipole moment

Monitoring the rotational dynamics of proteins at long timescales and hence over large changes in angle shows theoretically (Fig. 2c) and experimentally (Fig. 4c,e) that proteins with a dipole moment do not rotate randomly when they experience the  $MV m^{-1}$  electric field intensity inside the pore; instead, the proteins undergo biased Brownian rotation due to electric-field-induced torque on their dipole moment<sup>5</sup>. Quantifying this bias in orientation by fitting the intra-event  $\Delta I$  distribution from an individual resistive pulse with the convolution model made it possible to calculate a protein's dipole moment  $\mu$  by considering the potential energy landscape of a dipole in an electric field (Fig. 4b; and see Supplementary Notes 2, 6 and 9 and Supplementary Figs 25 and 26).

Figure 4e shows that the most probable values of dipole moment from this nanopore-based analysis agree well with expected values; the average deviation is less than 25%. With regard to the robustness of this method from experiment-to-experiment, the relative standard deviation of the most probable value from distributions of measured single-molecule values was 12% and compares well with dielectric impedance spectroscopy measurements<sup>7</sup>; however, the variation from event-to-event was large with a mean absolute deviation of 230% (Supplementary Fig. 25).

While the uncertainty in each stand-alone single-molecule measurement of dipole moment will have to be reduced in order to realize the full potential of this approach, this technique introduces the first experimental method for determining the dipole moment of individual proteins *in situ*. To this end, it exploits a fundamental advantage of single-molecule techniques, namely that statistical fluctuations of one particle are easier to interpret and to compare with theoretical models than it would be of an ensemble of particles. An additional advantage of this single-particle analysis is that it can estimate dipole moments in real time (Supplementary Fig. 10) and requires only pico- to nanomolar concentrations of proteins. In contrast, the standard method for measuring dipole moment, dielectric impedance spectroscopy, requires micromolar protein concentrations and significantly larger sample volumes<sup>7</sup>. We propose that dipole moments are excellent orthogonal protein descriptors since they are widely distributed among different proteins and not correlated with protein volume or charge (Supplementary Figs 17–19).

### Simulations confirm the feasibility of the approach

We confirmed the accuracy of these two approaches for multiparameter determination with simulated data that were generated from the theory of biased one-dimensional Brownian diffusion and convolved with current noise. Fitting either maximum  $\Delta I$  data from many simulated events or  $\Delta I(t)$  data from individual simulated events with the convolution model returned values of volume, shape, dipole moment and rotational diffusion coefficient that were in excellent agreement with the input parameters for the simulation (Supplementary Note 5 and Supplementary Figs 10–14). These purely theoretical results provide strong complementary evidence for the effectiveness of the analysis methods developed in this work.

### Multiparameter analysis improves protein classification

To assess the potential of nanopore-based identification and characterization of different proteins in a well-defined mixture, we repeated the characterization of glucose-6-phosphate dehydrogenase described in Fig. 3f and added a polyclonal anti-G6PDH IgG antibody. Thus, in the same experiment, single proteins of G6PDH and protein–protein complexes of G6PDH–IgG were passing through the nanopore. Analysis of intra-event  $\Delta I$  distributions from individual resistive pulses returned an estimate of the volume, shape, charge, rotational

diffusion coefficient and dipole moment for single particles passing through the pore. Figure 5 shows that this multiparameter–fingerprinting approach made it possible to distinguish G6PDH from the G6PDH–IgG complex by using a clustering algorithm to classify each translocation event (Fig. 5b; and see Supplementary Note 8 and Supplementary Fig. 30 for details)<sup>21</sup>. This analysis returned excellent estimates of the size and shape of G6PDH and the G6PDH–IgG complex (Fig. 5a,c). In contrast, employing the current standard practice of distinguishing proteins by the  $\Delta I$  values and translocation times of individual resistive pulses<sup>36</sup> underestimated the amount of the G6PDH–IgG complex formed by 90% and overestimated its volume by 70% (Supplementary Note 8). This analysis, therefore, provides proof of principle for nanopore-based characterization, identification and quantification at the single-protein level in a well-defined binary mixture and demonstrates the advantage of simultaneous multiparameter characterization for identifying individual proteins or protein–protein complexes over single-variate or bi-variate characterization.

### Conclusions

Based on the progress in nanopore-based DNA sequencing in the past 20 years<sup>3,37,38</sup>, we predict that improvements to the approach introduced here will further increase the potential of simultaneous, nanopore-based characterization of multiple protein descriptors on a single-molecule level<sup>39</sup>. For instance, the single-event (intra-event) analysis probably suffers from deviations in the pore geometry from a perfect cylinder. These irregularities, which are a consequence of the current state-of-the-art fabrication methods, affect the local resistance along the lumen of the pore and hence affect the precision with which the maximum and minimum  $\Delta I$  value can be determined. Novel fabrication methods such as He-ion beam fabrication produce pores that are almost perfectly cylindrical and should therefore minimize possible artifacts from this source of error<sup>40</sup>. In addition, the recent development of integrated CMOS current amplifiers<sup>41</sup>, which can be produced in parallel to record from hundreds of nanopores simultaneously while reaching at least ten-times higher bandwidth and three-times higher signal-to-noise ratio compared with the amplifier used in this work<sup>41</sup>, will increase the throughput and improve the precision and accuracy of determining the rotational dynamics of proteins on their journey through the pore. Such fast amplifiers may eliminate the need for tethering proteins to lipid anchors<sup>30</sup> while their improved signal-to-noise ratio combined with the recent development of low-noise nanopore chips<sup>42</sup> is likely to reduce the uncertainty in each determined parameter<sup>41,43</sup>. Furthermore, computational approaches that can model proteins with shapes more complex than simple spheroids<sup>44</sup> may increase the resolution of shape determination, while the capability to monitor current modulations with MHz bandwidths<sup>30,41</sup> may open up the possibility to follow transient changes in protein conformation and folding as well as to determine the shape of short-lived protein complexes whose structure and dynamics are not accessible by existing techniques.

We suggest that the ability to measure five parameters simultaneously on single proteins in real time, including parameters that can otherwise not be obtained on a single-molecule level, has transformative potential for the analysis and quantification of proteins as well as for the characterization of nanoparticles and nanoparticle assemblies. For instance, fast protein identification and quantification in complex mixtures is an unsolved problem<sup>2</sup>. Despite its tremendous capabilities, at present, mass spectrometry has currently limited throughput and is not broadly applicable to meet demand for routine protein analysis<sup>1,2</sup>. Two-dimensional gel electrophoresis remains one of the most important techniques for analysing complex protein samples, but its reproducibility is limited, and the method is slow and semi-quantitative. We propose that the work presented has the potential to replace methods such as 2D gel electrophoresis by providing additional protein descriptors, improved quantification and

reduced analysis time — although significant challenges will have to be addressed before resistive pulses from complex mixtures of small and large proteins can be interpreted. If these challenges can be met, then multiparameter characterization of single proteins in real time may ultimately reveal biochemically or clinically relevant static or dynamic heterogeneities (such as phosphorylations)<sup>45</sup>, allow monitoring an individual's proteome<sup>1</sup> and enable single-molecule sorting.

## Methods

Methods and any associated references are available in the [online version of the paper](#).

Received 9 November 2015; accepted 3 November 2016;  
published online 19 December 2016

## References

- Picotti, P. & Aebersold, R. Selected reaction monitoring-based proteomics: workflows, potential, pitfalls and future directions. *Nat. Methods* **9**, 555–566 (2012).
- Herr, A. E. Disruptive by design: a perspective on engineering in analytical chemistry. *Anal. Chem.* **85**, 7622–7628 (2013).
- Rusk, N. Disruptive proteomics. *Nat. Methods* **10**, 35–35 (2013).
- Sali, A., Glaeser, R., Earnest, T. & Baumeister, W. From words to literature in structural proteomics. *Nature* **422**, 216–225 (2003).
- Onclay, J. L. The investigation of proteins by dielectric measurements. *Chem. Rev.* **30**, 433–450 (1942).
- Felder, C. E., Prilusky, J., Silman, I. & Sussman, J. L. A server and database for dipole moments of proteins. *Nucl. Acids Res.* **35**, W512–W521 (2007).
- Singh, S., Yadav, S., Shire, S. & Kalonia, D. Dipole-dipole interaction in antibody solutions: correlation with viscosity behavior at high concentration. *Pharm. Res.* **31**, 2549–2558 (2014).
- Mehl, J. W., Onclay, J. L. & Simha, R. Viscosity and the shape of protein molecules. *Science* **92**, 132–133 (1940).
- Bonincontro, A. & Risuleo, G. Dielectric spectroscopy as a probe for the investigation of conformational properties of proteins. *Spectrochim. Acta Part A* **59**, 2677–2684 (2003).
- Howorka, S. & Siwy, Z. Nanopore analytics: sensing of single molecules. *Chem. Soc. Rev.* **38**, 2360–2384 (2009).
- Wei, R., Gatterdam, V., Wieneke, R., Tampe, R. & Rant, U. Stochastic sensing of proteins with receptor-modified solid-state nanopores. *Nat. Nanotech.* **7**, 257–263 (2012).
- Qin, Z. P., Zhe, J. A. & Wang, G. X. Effects of particle's off-axis position, shape, orientation and entry position on resistance changes of micro Coulter counting devices. *Meas. Sci. Technol.* **22**, 045804 (2011).
- Golibersuch, D. C. Observation of aspherical particle rotation in Poiseuille flow via the resistance pulse technique. Part 1. Application to human erythrocytes. *Biophys. J.* **13**, 265–280 (1973).
- Follega, D., Ledden, B., David, S. M. & Li, J. Electrical characterization of protein molecules by a solid-state nanopore. *Appl. Phys. Lett.* **91**, 053901 (2007).
- Sexton, L. T. *et al.* Resistive-pulse studies of proteins and protein/antibody complexes using a conical nanotube sensor. *J. Am. Chem. Soc.* **129**, 13144–13152 (2007).
- Yusko, E. C. *et al.* Controlling protein translocation through nanopores with bio-inspired fluid walls. *Nat. Nanotech.* **6**, 253–260 (2011).
- Robertson, J. W. F. *et al.* Single-molecule mass spectrometry in solution using a solitary nanopore. *Proc. Natl Acad. Sci. USA* **104**, 8207–8211 (2007).
- Raillon, C. *et al.* Nanopore detection of single molecule RNAP-DNA transcription complex. *Nano Lett.* **12**, 1157–1164 (2012).
- Soni, G. V. & Dekker, C. Detection of nucleosomal substructures using solid-state nanopores. *Nano Lett.* **12**, 3180–3186 (2012).
- Di Fiori, N. *et al.* Optoelectronic control of surface charge and translocation dynamics in solid-state nanopores. *Nat. Nanotech.* **8**, 946–951 (2013).
- Yusko, E. C. *et al.* Single-particle characterization of A $\beta$  oligomers in solution. *ACS Nano* **6**, 5909–5919 (2012).
- German, S. R., Hurd, T. S., White, H. S. & Mega, T. L. Sizing individual Au nanoparticles in solution with sub-nanometer resolution. *ACS Nano* **9**, 7186–7194 (2015).
- Nir, I., Huttner, D. & Meller, A. Direct sensing and discrimination among ubiquitin and ubiquitin chains using solid-state nanopores. *Biophys. J.* **108**, 2340–2349 (2015).
- Stefureac, R. I., Kachayev, A. & Lee, J. S. Modulation of the translocation of peptides through nanopores by the application of an AC electric field. *Chem. Commun.* **48**, 1928–1930 (2012).
- Li, J. *et al.* Ion-beam sculpting at nanometre length scales. *Nature* **412**, 166–169 (2001).
- DeBlois, R. W., Uzgiris, E. E., Cluxton, D. H. & Mazzone, H. M. Comparative measurements of size and polydispersity of several insect viruses. *Anal. Biochem.* **90**, 273–288 (1978).
- Smythe, W. R. Flow around a spheroid in a circular tube. *Phys. Fluids* **7**, 633–638 (1964).
- Fricke, H. The electric permittivity of a dilute suspension of membrane-covered ellipsoids. *J. Appl. Phys.* **24**, 644–646 (1953).
- Velick, S. & Gorin, M. The electrical conductance of suspensions of ellipsoids and its relation to the study of avian erythrocytes. *J. Gen. Physiol.* **23**, 753–771 (1940).
- Plesa, C. *et al.* Fast translocation of proteins through solid state nanopores. *Nano Lett.* **13**, 658–663 (2013).
- Hernandez-Ainsa, S. *et al.* Lipid-coated nanocapillaries for DNA sensing. *Analyst* **138**, 104–106 (2013).
- Woodside, M. T. *et al.* Direct measurement of the full, sequence-dependent folding landscape of a nucleic acid. *Science* **314**, 1001–1004 (2006).
- Deen, W. M. Hindered transport of large molecules in liquid-filled pores. *AIChE J.* **33**, 1409–1425 (1987).
- Happel, J. & Brenner, H. *Low Reynolds Number Hydrodynamics: With Special Applications to Particulate Media* 1st edn (Springer, 1983).
- Yuan, Y. & Axelrod, D. Subnanosecond polarized fluorescence photobleaching: rotational diffusion of acetylcholine receptors on developing muscle cells. *Biophys. J.* **69**, 690–700 (1995).
- Li, W. *et al.* Single protein molecule detection by glass nanopores. *ACS Nano* **7**, 4129–4134 (2013).
- Kasianowicz, J. J., Brandin, E., Branton, D. & Deamer, D. W. Characterization of individual polynucleotide molecules using a membrane channel. *Proc. Natl Acad. Sci. USA* **93**, 13770–13773 (1996).
- Manrao, E. A. *et al.* Reading DNA at single-nucleotide resolution with a mutant MspA nanopore and phi29 DNA polymerase. *Nat. Biotechnol.* **30**, 349–353 (2012).
- Nivala, J., Marks, D. B. & Akeson, M. Unfoldase-mediated protein translocation through an  $\alpha$ -hemolysin nanopore. *Nat. Biotechnol.* **31**, 247–250 (2013).
- Jijin, Y. *et al.* Rapid and precise scanning helium ion microscope milling of solid-state nanopores for biomolecule detection. *Nanotechnology* **22**, 285310 (2011).
- Rosenstein, J. K., Wanunu, M., Merchant, C. A., Drndic, M. & Shepard, K. L. Integrated nanopore sensing platform with sub-microsecond temporal resolution. *Nat. Methods* **9**, 487–492 (2012).
- Lee, M.-H. *et al.* A low-noise solid-state nanopore platform based on a highly insulating substrate. *Sci. Rep.* **4**, 7448 (2014).
- Uram, J. D., Ke, K. & Mayer, M. Noise and bandwidth of current recordings from submicrometer pores and nanopores. *ACS Nano* **2**, 857–872 (2008).
- Ortega, A., Amoros, D. & Garcia de la Torre, J. Prediction of hydrodynamic and other solution properties of rigid proteins from atomic- and residue-level models. *Biophys. J.* **101**, 892–898 (2011).
- De Pascalis, A. R. *et al.* Binding of ferredoxin to ferredoxin: NADP<sup>+</sup> oxidoreductase: the role of carboxyl groups, electrostatic surface potential, and molecular dipole moment. *Protein Sci.* **2**, 1126–1135 (1993).
- Pedone, D., Firmkes, M. & Rant, U. Data analysis of translocation events in nanopore experiments. *Anal. Chem.* **81**, 9689–9694 (2009).
- Talaga, D. S. & Li, J. L. Single-molecule protein unfolding in solid state nanopores. *J. Am. Chem. Soc.* **131**, 9287–9297 (2009).

## Acknowledgements

We thank C. L. Asbury for several helpful discussions and review of the manuscript. This work was supported by a Miller Faculty Scholar Award (M.M.), the Air Force Office of Scientific Research (M.M. and D.S., grant number FA9550-12-1-0435), Oxford Nanopore Technologies (M.M., grant number 350509-N016133), the National Institutes of Health (M.M., grant number 1R01GM081705), the National Human Genome Research Institute (J.L., grant numbers HG003290 and HG004776), J. Golovchenko's Harvard nanopore group for FIB pore preparation (J.L.), a Rackham Pre-Doctoral Fellowship from the University of Michigan (E.C.Y.), a Graduate Research Fellowship from the National Science Foundation (J.H.) and the Microfluidics in Biomedical Sciences Training Program from the NIH and BIBIB (B.R.B.).

## Author contributions

E.C.Y., B.R.B. and M.M. conceived and designed experiments, analysed data, and co-wrote the manuscript. E.C.Y., B.R.B., J.H. and O.M.E. performed nanopore experiments. R.C.R., N.C.W., S.N., A.R.H. and J.L. fabricated nanopores. M.P., D.S.K. and B.R.B. measured the dipole moments of proteins with impedance spectroscopy and provided constructive feedback on the manuscript. D.S. performed computational analyses of several protein crystal structures and provided guidance on statistical methods used in the manuscript.

## Additional information

Supplementary information is available in the [online version of the paper](#). Reprints and permissions information is available online at [www.nature.com/reprints](http://www.nature.com/reprints). Correspondence and requests for materials should be addressed to M.M.

## Competing financial interests

The authors declare no competing financial interests.

## Methods

**Materials.** All phospholipids were obtained from Avanti Polar Lipids. Bis (succinimidyl) penta(ethylene glycol) (21581) was purchased from Thermo Scientific. Monoclonal anti-biotin IgG<sub>1</sub> (B7653), polyclonal anti-G6PDH IgG, (A9521), GPI-anchored acetylcholinesterase (C0663), glucose-6-phosphate dehydrogenase (G5885), L-lactate dehydrogenase (59747), bovine serum albumin (A7638),  $\alpha$ -amylase (A4551) and streptavidin were purchased from Sigma Aldrich. Polyclonal anti-biotin IgG-Fab fragments (800-101-098) were purchased from Rockland and  $\beta$ -phycoerythrin (P-800) was purchased from Life Technologies.

**Methods of nanopore-based sensing experiments.** To sense proteins, we first formed a supported lipid bilayer containing either 0.15 mol% 1,2-dipalmitoyl-*sn*-glycero-3-phosphoethanolamine-*N*-capbiotinyl (biotin-PE) lipids or 1 mol% 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphoethanolamine (POPE) lipid in a background of 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine (POPC) lipids (Avanti Polar Lipids). We described details of the bilayer formation in ref. 16. The dimensions of all nanopores are shown in Supplementary Fig. 31. When biotin-PE lipids were present in the bilayer, we added a solution containing anti-biotin IgG<sub>1</sub>, Fab, or GPI-anchored acetylcholinesterase to the top solution compartment of the fluidic set-up such that the final concentration of protein ranged from 5 pM to 10 nM. When sensing GPI-anchored acetylcholinesterase, we started recording resistive pulses after incubating the bilayer-coated nanopore for 1 h with GPI-anchored acetylcholinesterase (where the solution was 150 mM KCl, 10 mM HEPES, pH = 7.4) to allow time for the GPI-lipid anchor of the protein to insert into the fluid lipid bilayer coating. When POPE lipids were present in the bilayer, we first dissolved bis(succinimidyl) penta(ethylene glycol), a bifunctional crosslinker, in a buffer containing 2 M KCl and 100 mM KHCO<sub>3</sub> (pH = 8.4) and immediately added this solution to the top compartment of the fluidic set-up such that the final concentration of crosslinker was 10 mg ml<sup>-1</sup>. After 10 min, we rinsed away excess crosslinker and subsequently added  $\beta$ -phycoerythrin, glucose-6-phosphate dehydrogenase, L-lactate dehydrogenase, bovine serum albumin,  $\alpha$ -amylase, or butyrylcholinesterase dissolved in the same buffer as the preceding step to the top compartment such that final protein concentration ranged from 1 to 3  $\mu$ M. After at least 30 min, we rinsed away excess protein and began recording. Therefore, depending on the interaction between the protein of interest and its lipid anchor, the limit of detection of this method ranged from low picomolar to low micromolar concentrations, corresponding to a total protein amount of 10<sup>-16</sup> to 10<sup>-10</sup> moles; these minimal concentrations were required to observe a significant frequency of translocation events. We recorded resistive pulses at an applied potential difference of -0.04 to -0.115 V with the polarity referring to the top fluid compartment relative to the bottom fluid compartment, which was connected to ground. The electrolyte contained 2 M KCl with either: 10 mM HEPES at pH 6.5 for experiments with

GPI-anchored acetylcholinesterase; 10 mM HEPES at pH 7.4 for experiments with IgG, Fab,  $\alpha$ -amylase, butyrylcholinesterase, and streptavidin; 10 mM citrate (C<sub>6</sub>H<sub>7</sub>KO<sub>7</sub>) at pH 5.1 for experiments with  $\beta$ -phycoerythrin; 10 mM C<sub>6</sub>H<sub>7</sub>KO<sub>7</sub> at pH 5.2 for experiments with bovine serum albumin; or 10 mM C<sub>6</sub>H<sub>7</sub>KO<sub>7</sub> at pH 6.1 for experiments with glucose-6-phosphate dehydrogenase and L-lactate dehydrogenase. We used Ag/AgCl pellet electrodes (Warner Instruments) to monitor ionic currents through electrolyte-filled nanopores with a patch-clamp amplifier (Axopatch 200B, Molecular Devices) in voltage-clamp mode (that is, at constant applied voltage). We set the analog low-pass filter of the amplifier to a cutoff frequency of 100 kHz. We used a digitizer (Digidata 1322) with a sampling frequency of 500 kHz in combination with a program written in LabView to acquire and store data<sup>43</sup>. To distinguish resistive pulses reliably from the electrical noise, we first filtered the data digitally with a Gaussian low-pass filter ( $f_c = 15$  kHz) in MATLAB and then used a modified form of the custom-written MATLAB routine described in ref. 46. We defined the start of a resistive pulse by a resistive decrease in the magnitude of the current past a threshold value that we set to 5 $\times$  the standard deviation of the noise of the baseline current. The subsequent return of the current past a second threshold, which we set to one standard deviation of the noise in the baseline current, and towards the baseline value, marked the end of the resistive pulse<sup>16</sup>. We calculated the translocation time,  $t_d$ , as the width of individual resistive-pulse at half of their peak amplitude, also known as the full-width at half-maximum value<sup>16,47</sup>. From this analysis, we obtained the  $\Delta I$  and  $t_d$  values for each resistive pulse, and we only analysed  $\Delta I$  values for resistive-pulses with  $t_d$  values greater than 50  $\mu$ s, since resistive pulses with translocation times shorter than 50  $\mu$ s have attenuated  $\Delta I$  values due to the low-pass filter<sup>16,46</sup>. Refer to Supplementary Figs S19 and S20 (from ref. 16) for a detailed characterization of the frequency response of the recording set-up used here.

With regard to the success rate of the experiments reported here, we used a total of 68 different nanopores for this work and 21 of these nanopores (31%) yielded measurements. Experiments generally failed due to one of three reasons: (1) the baseline current was lower than expected based on pore geometry and electrolyte conductivity prior to coating the nanopore with a lipid bilayer ( $I_{\text{baseline}} < 0.9 \times I_{\text{expected}}$ ); (2) the nanopore did not coat; or (3) the baseline current after coating was too noisy to detect translocation events. For the 68 nanopores, we obtained the expected baseline current in 73% of attempts, successfully coated the pore in 37% of attempts (cumulative success rate = 27%), and achieved sufficiently low noise for recording after successfully coating the pore in 46% of attempts (cumulative success rate = 12%). These statistics indicate that approximately 1 in 10 experiments yielded a measurement, on average. The success rate was, however, highly dependent on the nanopore chip being used: a subset of approximately 10 nanopores were coated successfully in ~80% of attempts until they abruptly failed irreversibly at the first stage described above; on average this failure occurred after 16 experiments.