





A generalizable nanopore sensor for highly specific protein detection at single-molecule precision

Received: 24 September 2022

Accepted: 23 February 2023

Published online: 20 March 2023

 Check for updates

Mohammad Ahmad¹, Jeung-Hoi Ha², Lauren A. Mayse ^{1,3}, Maria F. Presti², Aaron J. Wolfe^{1,4,5,6}, Kelsey J. Moody^{1,4,5,6}, Stewart N. Loh ² & Liviu Movileanu ^{1,3,7} 

Protein detection has wide-ranging implications in molecular diagnostics. Substantial progress has been made in protein analytics using nanopores and the resistive-pulse technique. Yet, a long-standing challenge is implementing specific interfaces for detecting proteins without the steric hindrance of the pore interior. Here, we formulate a class of sensing elements made of a programmable antibody-mimetic binder fused to a monomeric protein nanopore. This way, such a modular design significantly expands the utility of nanopore sensors to numerous proteins while preserving their architecture, specificity, and sensitivity. We prove the power of this approach by developing and validating nanopore sensors for protein analytes that drastically vary in size, charge, and structural complexity. These analytes produce unique electrical signatures that depend on their identity and quantity and the binder-analyte assembly at the nanopore tip. The outcomes of this work could impact biomedical diagnostics by providing a fundamental basis for biomarker detection in biofluids.

Identifying and quantifying protein biomarkers is a pressing demand in precision and personalized medicine¹. Recent advancements in functional proteomics indicate that there are yet numerous unexplored proteins with potential consequences for the progression of pathological conditions². For example, the development of various hematological malignancies and solid tumors is associated with changes in specific protein expressions. In general, protein biomarkers are upregulated under oncogenic conditions. Clinical examinations of such biomarkers are instrumental in prognostics, diagnostics, and therapeutics. A persistent challenge in protein detection in biofluids is the occurrence of nonspecific bindings of assay reagents with immobilization surfaces and protein recognition elements, amplifying the

background noise. In addition, existing techniques cannot be utilized for a wide dynamic range because of deterioration in the signal-to-noise ratio at low analyte levels. Therefore, there is an increasing necessity to create highly specific and sensitive protein-sensing approaches that employ rapid signal responses to various biochemical stimuli^{3,4}.

Molecular details of protein detection are illuminated using single-molecule methods^{3–5}. In particular, single-molecule sensing with nanopores^{6–8} using the resistive-pulse technique⁹ is adaptable to parallel recording technologies¹⁰. In the past decade, substantial progress has been made in peptide and protein analytics using nanopore sensors^{11–13}. Nowadays, nanopores are fabricated using a broad range

¹Department of Physics, Syracuse University, 201 Physics Building, Syracuse, NY 13244-1130, USA. ²Department of Biochemistry and Molecular Biology, State University of New York-Upstate Medical University, 4249 Weiskotten Hall, 766 Irving Avenue, Syracuse, NY 13210, USA. ³Department of Biomedical and Chemical Engineering, Syracuse University, 329 Link Hall, Syracuse, NY 13244, USA. ⁴Ichor Life Sciences, Inc., 2561 US Route 11, LaFayette, NY 13084, USA. ⁵Lewis School of Health Sciences, Clarkson University, 8 Clarkson Avenue, Potsdam, NY 13699, USA. ⁶Department of Chemistry, College of Environmental Science and Forestry, State University of New York, 1 Forestry Drive, Syracuse, NY 13210, USA. ⁷The BioInspired Institute, Syracuse University, Syracuse, NY 13244, USA. ✉e-mail: lmovilea@syr.edu

of scaffolds, either in organic or inorganic materials^{13,14}. These sensing elements can identify and quantify peptides at single-amino acid resolution^{15,16}. Moreover, nanopores represent ultrasensitive sensors for revealing critical features of proteins, such as shape and size¹⁷, post-translational modifications^{18,19}, enzymatic activity^{20,21}, and mechanical unfolding^{22,23}. In addition, nanopores are utilized to fingerprint proteins using enzymatic degradation^{24,25}. In recent years, these sensors were engineered and formulated for challenging aspects in single-molecule protein sequencing^{26,27}. Despite many benefits, this approach requires the targeted proteins to partition into the nanopore interior. Hence, the detection is conducted under steric restrictions of the nanopore confinement, potentially impairing the strength of specific interactions. If the nanopore diameter is smaller than the hydrodynamic radius of the protein analyte, then this is another obstacle preventing its detection in the natively folded state.

Detecting single proteins outside the nanopore is a practical alternative to sampling the complexity of protein recognition events^{28–30}. This task would necessitate an external protein binder (e.g., receptor) covalently attached to a nanopore. However, a transducing mechanism is needed to convert the physical captures and releases of a protein analyte (e.g., its ligand) into a specific electrical signature of the sensor. In addition, changing the system to a different binder-analyte pair requires a lengthy and tedious optimization process that includes amplified difficulties. The heterogeneous architecture, size, charge, and other traits of different binders need extensive protein engineering. This prerequisite is critical for each sensor for a given protein analyte. Earlier studies have suggested that these protein sensors may be limited to established protein fragments of ~100 residues^{31,32}. For example, large protein binders likely induce additional steric constraints, precluding the clearance of the space around the pore opening. Moreover, the interaction interface of the binder must be fully accessible to the protein analyte.

Here, we propose a class of sensing elements for probing proteins at a single-detector precision. These sensors will have an antibody-mimetic protein binder engineered on the tFhuA nanopore³¹, a monomeric β -barrel scaffold, via a flexible tether. Protein recognition elements, such as proteins^{31,32}, or peptides³³, can be covalently tethered at the N terminus of tFhuA through a flexible (GGG)₂ peptide tether without the deterioration of its membrane-embedded structure and pore-forming features. Furthermore, this hexapeptide maintains the protein binder in the proximity of the pore opening to acquire a current modulation upon individual binding events. In this study, we apply a similar strategy to develop nanopore sensors as single-polypeptide chain proteins. They are refolded in detergents and reconstituted into a synthetic lipid bilayer at a single-molecule level (Fig. 1). Hence, all data presented in this article are determined at single-molecule precision. We demonstrate that by changing only the binding interface, a different binder-containing nanopore sensor can be obtained and readily implemented into detecting a specific protein. This strategy maintains the sensor's architecture, high sensitivity, and specificity while featuring its generalization to numerous protein analytes.

In this article, the binder is a monobody, a recombinant protein based on the 94-residue fibronectin type-III (FN3) domain³⁴. Target-specific monobodies are preferable over other antibody-mimetic scaffolds for the following reasons: (i) the monobody has a relatively smaller hydrodynamic radius. This characteristic maintains the tFhuA pore partially open, so it enables a readable output signal through the passage of ions; (ii) the monobody does not partition into the tFhuA pore lumen, and its binding interface is exposed to the aqueous phase for a specific protein target; (iii) monobodies, unlike nanobodies, lack disulfide bridges³⁵. Therefore, they provide an opportunity for straightforward expression, purification, and refolding procedures; and (iv) monobodies can be selected from a broad range of variants against numerous target proteins³⁶. Using monobody-based nanopore

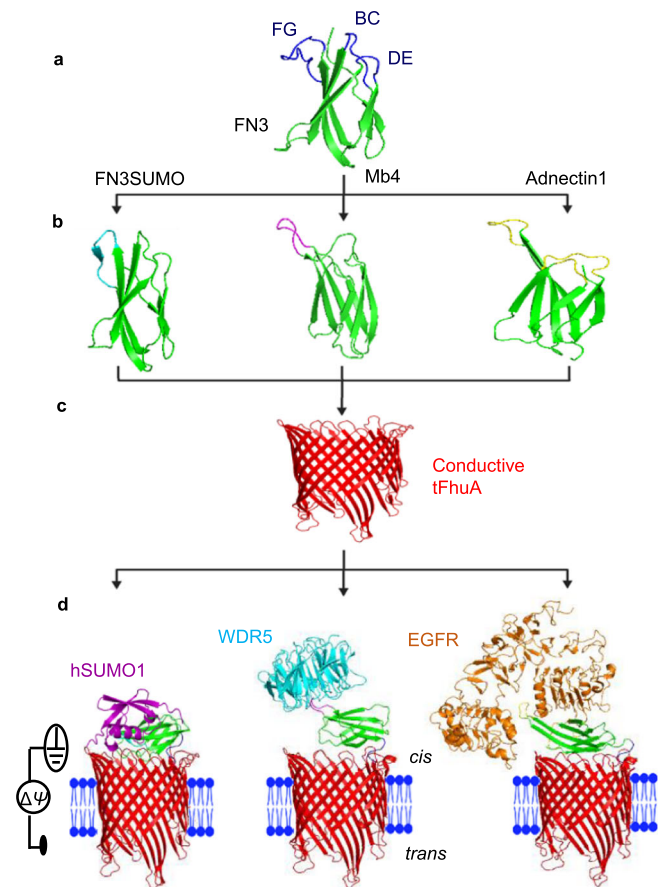


Fig. 1 | Rational protein design of a generalizable nanopore sensor for protein detection. **a** A tenth fibronectin type-III domain (FN3)³⁴ or monobody (in green) with the FG, BC, and DE loops highlighted in blue. **b** The FN3 variants, where cyan, magenta, and yellow were used to represent the binding loops in FN3SUMO⁴⁵, Mb4⁴⁶, and Adnectin1⁴⁷, respectively. **c** tFhuA (in red)³¹, a monomeric β barrel with average internal diameters of ~2.6 and ~3.9 nm, as measured from side chain to side chain. **d** Functional reconstitution of nanopore sensors into a lipid bilayer. The hSUMO1-specific nanopore is a single-polypeptide unit that comprises a tFhuA, a (GGG)₂ tether, and an FN3SUMO monobody (left). WDR5- and EGFR-binding monobodies, Mb4 (center) and Adnectin1 (right), respectively, were also fused to tFhuA in the same way as FN3SUMO. hSUMO1, WDR5, and EGFR are marked in magenta, cyan, and brown, respectively. $\Delta\psi$ indicates the applied transmembrane potential between the *cis* and *trans* sides of the membrane. The monobody-analyte complexes are shown as well. The structures of all sensors were predicted by AlphaFold2^{34,35}.

sensors with varying binding interfaces, it is possible to detect different proteins that vary substantially in their structural and functional properties. When subjected to a biofluid, this class of sensing elements can report the presence of a protein biomarker at a single-molecule level. This tactic will not only enable overcoming the abovementioned challenges but will also motivate the widespread applications of these sensors.

Here, we develop monobody-based sensors for three targeted analytes: (i) human small ubiquitin-related modifier 1 (hSUMO1), a model protein with implications in various cellular processes, such as DNA damage repair, chromosome dynamics, and cell cycle^{37,38}; (ii) WD40 repeat protein 5 (WDR5)³⁹, a chromatin-associated protein hub involved in the epigenetic regulation of histone 3 lysine 4 (H3K4) methylation. WDR5 is a putative biomarker because it is overexpressed under various oncogenic conditions^{40,41}; and (iii) epidermal growth factor receptor (EGFR), a prognosis protein biomarker in lung, colorectal, and breast cancers^{42–44} (Fig. 1; Supplementary Table 1). Therefore, we create three sensors using FN3SUMO⁴⁵, Mb4⁴⁶, and

Adnectin1⁴⁷ monobodies as binders against hSUMO1, WDR5, and the ectodomain of EGFR, respectively. These monobody-based sensors are denoted by FN3SUMO-tFhuA, Mb4-tFhuA, and Adnectin1-tFhuA, respectively (Supplementary Fig. 1).

Results and discussion

Development of monobody-based nanopore sensors

We connected three distinct monobodies to the N-terminus of tFhuA using our standard linker, (GGG)₂. We employed AlphaFold2 to predict the overall three-dimensional conformation of the monobody-tFhuA fusion proteins^{48,49}. Hence, our computational prediction of nanopore sensors allowed us to visualize the orientation of monobodies in these structures. We observed that the predicted Local Distance Difference Test (pLDDT), a confidence score for each residue, was between 80 and 100 for most residues of the FN3-tFhuA (Supplementary Fig. 2a, b). This model illustrates that FN3 orients almost perpendicularly on the central axis of tFhuA (Supplementary Fig. 2c). This finding is likely due to long-range electrostatic interactions between clusters of negative charges on tFhuA β turns and positive charges on FN3 loops (Supplementary Fig. 3). Similar results were obtained with FN3SUMO-tFhuA, Mb4-tFhuA, and Adnectin1-tFhuA (Supplementary Fig. 4). Therefore, FN3 monobodies in all sensors potentially block a substantial ionic flow through tFhuA. Inspecting all sensors at a transmembrane potential of +40 mV revealed a relatively quiet single-channel electrical current recorded with FN3SUMO-tFhuA and Mb4-tFhuA, and a slightly noisy signal acquired with Adnectin1-tFhuA (Supplementary Figs. 5 and 6). The unitary conductance of FN3SUMO-tFhuA, Mb4-tFhuA, and Adnectin1-tFhuA were (mean \pm s.d.) 0.81 ± 0.03 nS, 0.99 ± 0.04 nS, and 0.90 ± 0.02 nS (Supplementary Table 2), respectively. These are significant conductance reductions compared to the unmodified tFhuA (1.5 ± 0.1 nS) (Supplementary Fig. 7)³³. This finding is in accord with the predictions made by AlphaFold2.

Real-time and label-free detection of hSUMO1 using a single FN3SUMO-tFhuA nanopore sensor

A single FN3SUMO-tFhuA sensor was functionally reconstituted into a lipid membrane at an applied transmembrane potential of +40 mV. The presence of hSUMO1 in the *cis* compartment at nanomolar concentrations produced frequent current blockades (Fig. 2a and Supplementary Fig. 8) between O_{on} open-substate and O_{off} closed-substate. Their normalized current amplitude, A/I_0 , was $(91.5 \pm 0.7)\%$. Here, I_0 and A denote the single-channel current of the hSUMO1-released substate and the current amplitude of hSUMO1-produced current blockades, respectively (Fig. 2b, c). In addition, infrequent and brief current spikes were observed when hSUMO1 was added to the *cis* side of an unmodified tFhuA-containing bilayer (Supplementary Fig. 9). Taken together, these negative-control measurements indicate that hSUMO1 did not produce any significant current blockades due to nonspecific interactions with the *cis* opening of the nanopore.

Moreover, hSUMO1-captured events were noted as concentration-dependent (Fig. 2a). hSUMO1-released and hSUMO1-captured events recorded with FN3SUMO-tFhuA corresponded to the open-substate, O_{on} , and closed-substate, O_{off} , respectively. However, hSUMO1-captured events were not detectable when hSUMO1 was added to the *trans* compartment (Supplementary Fig. 10), confirming that tFhuA and its derivatives insert into the membrane with a single orientation⁵⁰.

Next, we pursued detailed statistical analyses of the hSUMO1-released, and hSUMO1-captured durations, whose mean values were denoted by τ_{on} and τ_{off} , respectively. The maximum likelihood method⁵¹ and logarithm likelihood ratio (LLR) tests⁵² were employed to determine the distribution model of these time constants. The duration of hSUMO1-released and hSUMO1-captured events showed a single-exponential distribution in the form of a single-peak in a semi-logarithmic plot (Fig. 2d,e). Although the bin size was identical in these

histograms, we represented them differently for clarity. Increasing the hSUMO1 concentration, [hSUMO1], decreased the τ_{on} but did not alter τ_{off} (Supplementary Table 3). The association rate constants, k_{on} , were consistent for all [hSUMO1] values (Supplementary Table 4). Here, $k_{on} = 1/([\text{hSUMO1}] \tau_{on})$. In addition, the frequency of hSUMO1-captured events, f , where $f = 1/\tau_{on}$, was proportional to [hSUMO1] in a ratio 1:1 (Fig. 2f), indicating a bimolecular association process of the hSUMO1-FN3SUMO complex. Using the linear fit of $f([\text{hSUMO1}])$, we obtained a k_{on} value (mean \pm s.e.m.) of $(1.12 \pm 0.02) \times 10^8 \text{ M}^{-1} \text{ s}^{-1}$. The dissociation rate constant k_{off} was determined as the reciprocal of the mean hSUMO1-captured durations ($1/\tau_{off}$). This value was independent of [hSUMO1] (Fig. 2g), suggesting a unimolecular dissociation mechanism of the hSUMO1-FN3SUMO complex. A linear fit of $k_{off}([\text{hSUMO1}])$ versus [hSUMO1] resulted in its mean \pm s.e.m. of $74.5 \pm 2.4 \text{ s}^{-1}$, to yield an equilibrium dissociation constant (K_D) of $665 \pm 24 \text{ nM}$ (Supplementary Table 5).

Detection of a chromatin-associated protein hub using a single Mb4-tFhuA nanopore sensor

We employed the same approach and experimental conditions to detect WDR5 using a functionally reconstituted Mb4-tFhuA sensor into a lipid bilayer. When added to the *cis* compartment at nanomolar concentrations, WDR5 produced frequent current blockades (Fig. 3a; Supplementary Fig. 11) between O_{on} open-substate and O_{off} partly closed-substate with a normalized current amplitude of $(14 \pm 1)\%$ (Fig. 3b, c). Again, this kind of current blockades was not noted when an unmodified tFhuA was exposed to WDR5 added to the *cis* side (Supplementary Fig. 12) or when Mb4-tFhuA was subjected to WDR5 added to the *trans* side (Supplementary Fig. 13). These findings suggest that specific WDR5-Mb4 interactions bring about WDR5-induced current blockades. WDR5-released (O_{on}) and WDR5-captured (O_{off}) events also followed a single-exponential distribution (Fig. 3d, e). In addition, the frequency of WDR5-captured events was proportional to the WDR5 concentration, [WDR5] (Fig. 3f), whereas their duration was independent of [WDR5] (Fig. 3e, g; Supplementary Tables 6 and 7). Using linear fits of the functions $f([\text{WDR5}])$ and $k_{off}([\text{WDR5}])$, we obtained a k_{on} value (mean \pm s.e.m.) of $(0.83 \pm 0.01) \times 10^8 \text{ M}^{-1} \text{ s}^{-1}$ and a k_{off} value (mean \pm s.e.m.) of $72.4 \pm 3.7 \text{ s}^{-1}$, resulting a K_D of $872 \pm 45 \text{ nM}$ (Supplementary Table 8). It should be noted the kinetics of WDR5-Mb4 interactions undergo fast association and dissociation rates, which were also detected with hSUMO1-FN3SUMO interactions.

An orthogonal method proves the rapid association and dissociation kinetics of WDR5-Mb4 interactions

To validate the fast kinetics recorded with the Mb4-tFhuA sensor, we performed additional measurements using biolayer interferometry (BLI)⁵³. Mb4-tFhuA-containing micelles were immobilized onto the BLI sensor surface via a cysteine sulfhydryl engineered on the external L4 loop of tFhuA for biotin-streptavidin chemistry (Methods; Supplementary Fig. 14a). Hence, this experimental design mimics a sensing measurement with an Mb4-tFhuA sensor reconstituted into a lipid bilayer. WDR5 was added to different wells at increased concentrations. The association phases were recorded in real-time by placing the BLI sensors in WDR5-containing wells (Supplementary Fig. 14b). The dissociation phases were then recorded by putting the same BLI sensors in WDR5-free wells. However, the rates of these kinetics are beyond the time resolution of BLI. Nevertheless, BLI sensorgrams acquired at various WDR5 concentrations qualitatively confirm the rapid kinetics of association and dissociation of WDR5-Mb4 interactions noted with the Mb4-tFhuA sensor (Supplementary Table 8).

A single Adnectin1-tFhuA nanopore sensor reveals bimodal protein recognition of EGFR

The ectodomain of EGFR is proteolytically released into the bloodstream, allowing this biomarker to be used for screening, diagnosis,

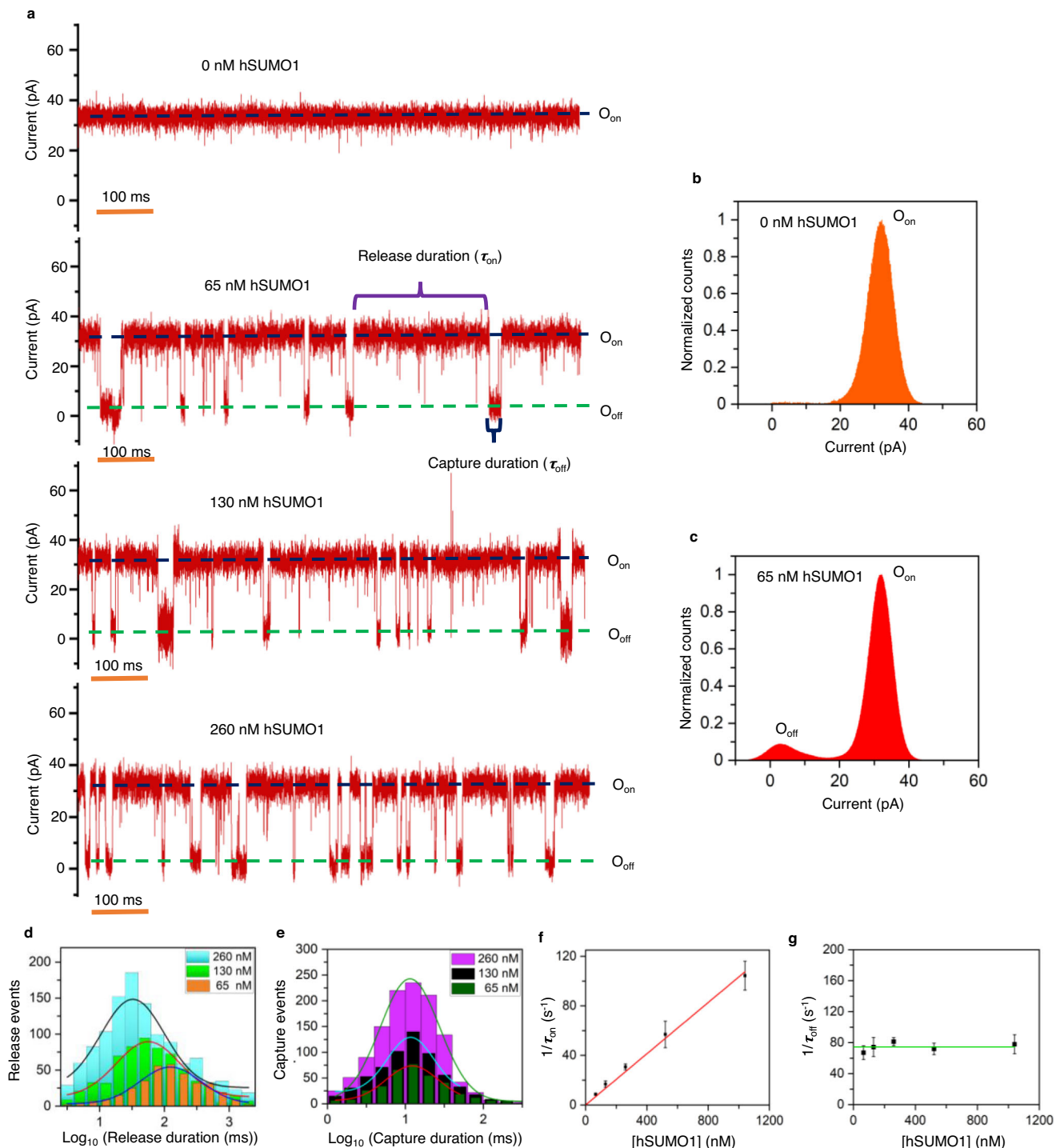


Fig. 2 | Real-time and label-free detection of hSUMO1. **a** Representative single-channel electrical traces (in red) of FN3SUMO-tFhuA in the presence of 0, 65, 130, and 260 nM hSUMO1. O_{on} and O_{off} are the hSUMO1-released and hSUMO1-captured substates, respectively. τ_{on} and τ_{off} are the durations of the hSUMO1-released and hSUMO1-captured events, respectively. These single-channel electrical signatures were replicated in $n = 3$ independent experiments. The applied transmembrane potential was +40 mV. Single-channel electrical traces were further low-pass filtered at 3 kHz using an 8-pole Bessel filter. **b** A current histogram (in orange) of the O_{on} substate of FN3SUMO-tFhuA. The current amplitude (mean \pm s.e.m.) of the O_{on} substate was 32.1 ± 0.1 pA. **c** A current histogram (in red) of the O_{on} and O_{off} substates of FN3SUMO-tFhuA at 65 nM hSUMO1. The current amplitude (mean \pm s.e.m.) of the O_{off} substate was 2.9 ± 0.1 pA. **d** Histograms of τ_{on} at various hSUMO1

concentrations, [hSUMO1]. τ_{on} (mean \pm s.e.m.) were 125 ± 4 ms (number of events: $N = 349$), 69 ± 5 ms ($N = 623$), and 35 ± 1 ms ($N = 1168$) at [hSUMO1] values of 65 nM, 130 nM, and 260 nM, respectively. **e** Histograms of τ_{off} at various [hSUMO1] values. τ_{off} (mean \pm s.e.m.) were 15 ± 1 ms ($N = 354$ events), 16 ± 1 ms ($N = 633$), and 14 ± 1 ms ($N = 1180$) at [hSUMO1] values of 65 nM, 130 nM, and 260 nM, respectively. **f** Dependence of the event frequency in the form of $1/\tau_{on}$ on [hSUMO1]. The slope of the linear fit (in red) of $1/\tau_{on}$ versus [hSUMO1] is the association rate constant, k_{on} , of hSUMO1-FN3SUMO interactions because $k_{on} = 1/(\tau_{on}[\text{hSUMO1}])$. **g** Dependence of $1/\tau_{off}$ on [hSUMO1]. The horizontal line is an average fit (in green) of the $(1/\tau_{off})$ data points recorded for various [hSUMO1] values. Data points in panels **f** and **g** represent mean \pm s.d. obtained from $n = 3$ different experiments. Source data are provided as a Source Data file.

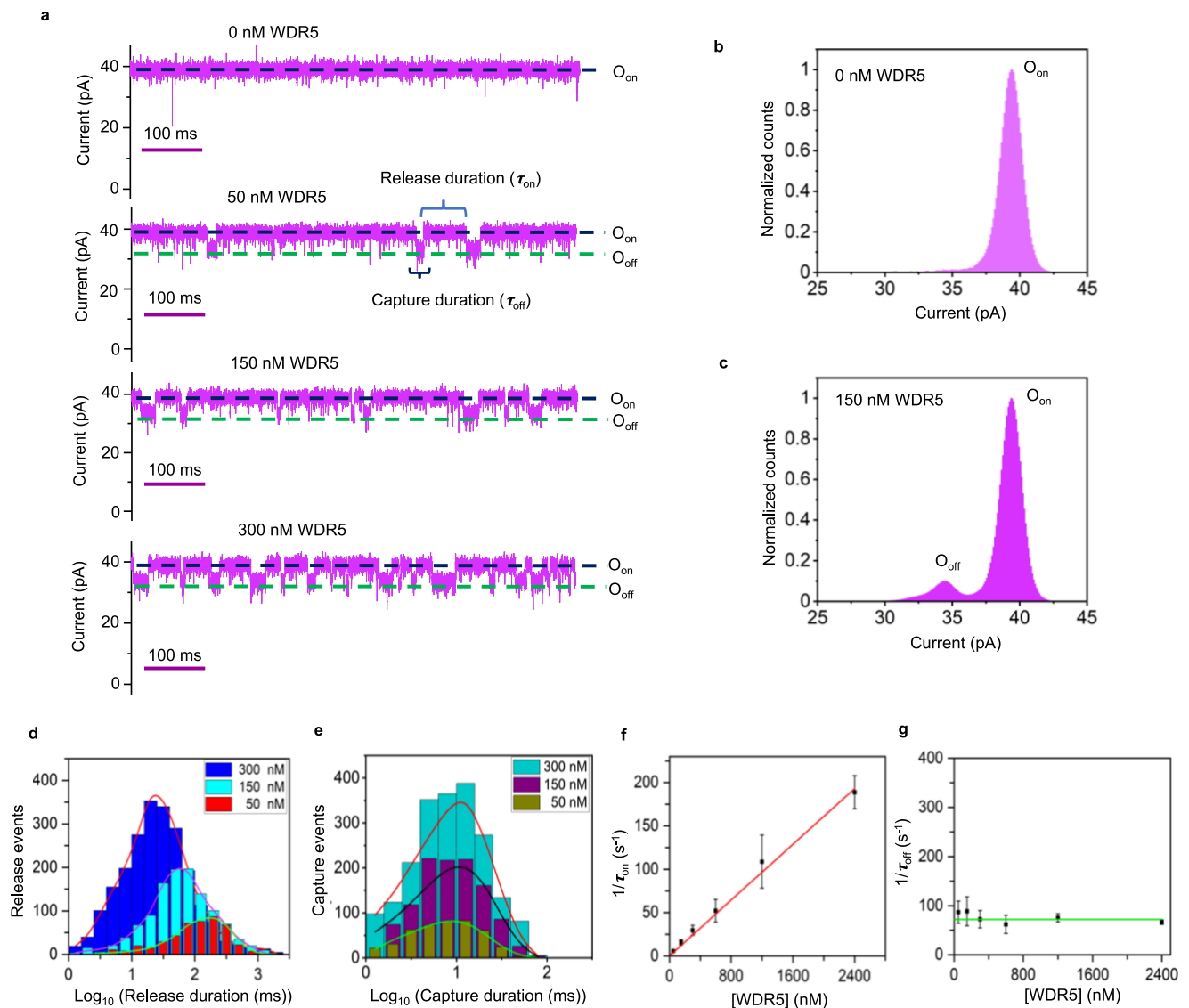


Fig. 3 | Single-molecule sensing of WDR5. **a** Representative single-channel electrical traces (in magenta) of Mb4-tFhuA in the presence of 0, 50, 150, and 300 nM WDR5. O_{on} and O_{off} are the WDR5-released and WDR5-captured substates, respectively. τ_{on} and τ_{off} are the durations of the WDR5-released and WDR5-captured events, respectively. These single-channel electrical signatures were replicated in $n = 3$ independent experiments. The other conditions were the same as those in Fig. 2. **b** A current histogram (in magenta) of the O_{on} substate of Mb4-tFhuA. The current (mean \pm s.e.m.) corresponding to the O_{on} substate was 39.5 ± 0.1 pA. **c** A current histogram (in magenta) of the O_{on} and O_{off} substates of Mb4-tFhuA at 150 nM WDR5. The current (mean \pm s.e.m.) corresponding to the O_{off} substate was 34.5 ± 0.1 pA. **d** Histograms of the WDR5-released durations (τ_{on}) at

various WDR5 concentrations, [WDR5]. τ_{on} (mean \pm s.e.m.) were 178 ± 6 ms (number of events: $N = 466$), 65 ± 7 ms ($N = 1175$), and 34 ± 4 ms ($N = 2235$) at [WDR5] values of 50 nM, 150 nM, and 300 nM, respectively. **e** Histograms of the WDR5-captured durations (τ_{off}) at various [WDR5] values. τ_{off} (mean \pm s.e.m.) were 12 ± 2 ms ($N = 441$ events), 10 ± 3 ms ($N = 1127$), and 14 ± 2 ms ($N = 2034$) at [WDR5] values of 50 nM, 150 nM, and 300 nM, respectively. **f** Plot (in red) illustrating the dependence of the event frequency in the form of $1/\tau_{on}$ on [WDR5]. **g** Plot (in green) illustrating the dependence of $1/\tau_{off}$ on [WDR5]. The horizontal line is an average fit of the $(1/\tau_{off})$ data points. Data points in panels **f** and **g** represent mean \pm s.d. obtained from $n = 3$ different experiments. Source data are provided as a Source Data file.

and disease progression⁵⁴. Hence, we employed Adnectin-1 against the ectodomain of EGFR⁴⁷. A single Adnectin1-tFhuA sensor exhibited some current noise at +40 mV (Supplementary Fig. 4c). However, its traces showed a relatively quiet signature at a lower transmembrane potential of +20 mV (Fig. 4a and Supplementary Fig. 15). Interestingly, when EGFR was added to the *cis* side of the bilayer containing the Adnectin1-tFhuA sensor, reversible current blockades were observed in a broad temporal range and with various current amplitudes. In contrast, we noted only low-amplitude and brief current spikes when EGFR was added to the *cis* side of the bilayer containing tFhuA alone (Supplementary Fig. 16). A two-peak distribution was found for the current amplitudes of individual EGFR-captured events (Fig. 4b). For example, at 40 nM EGFR, the normalized current blockades of the two

peaks were $(65.0 \pm 2.1)\%$ and $(86.1 \pm 1.6)\%$ with the probabilities of 0.72 ± 0.02 and 0.28 ± 0.02 , respectively. Furthermore, the relative position and probability of these peaks were independent of the EGFR concentration, [EGFR] (Supplementary Table 9).

EGFR-released (τ_{on}) and EGFR-captured (τ_{off}) durations followed single-peak and double-peak event distributions (Fig. 4c, d; Supplementary Tables 10–13), respectively, as judged by the maximum likelihood method⁵¹ and logarithm likelihood ratio (LLR) tests⁵². Hence, our statistical analyses revealed two subpopulations of binding events, the short-lived and long-lived EGFR-captured events, whose durations were $\tau_{off-1} = -80$ ms and $\tau_{off-2} = -1$ s, respectively. Interestingly, the probabilities of short-lived EGFR capture durations, P_1 , were close to those of low-amplitude current blockades (Supplementary Tables 9

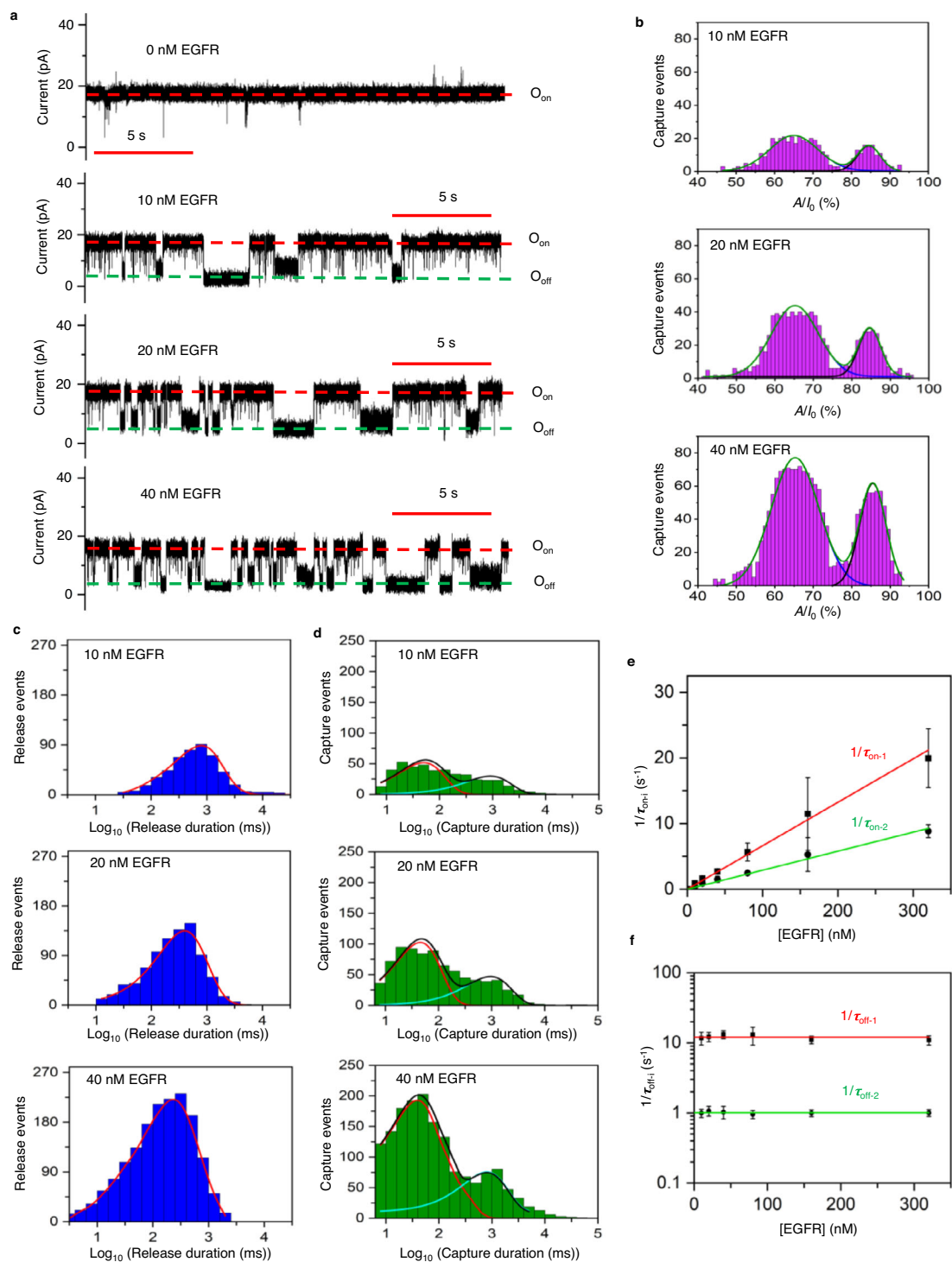


Fig. 4 | EGFR exhibits a bimodal protein recognition. **a** Single-channel traces (in black) of Adnectin1-tFluA in the presence of EGFR, which were filtered at 2 kHz. O_{on} and O_{off} are the EGFR-released and EGFR-captured substates, respectively. These traces were replicated in $n = 3$ independent experiments. The transmembrane potential was +20 mV. **b** Event histograms (in magenta) of the normalized current blockades A/I_0 , where A and I_0 are the amplitude of blockades and the amplitude of the O_{on} substate, respectively. The cumulative fits are marked in green. The blue and black curves indicate fits of low- and large-amplitude blockades, respectively. For 10 nM EGFR, these values (mean \pm s.e.m.) were $(65.0 \pm 0.3)\%$ and $(84.5 \pm 0.3)\%$, respectively (number of events, $N = 467$). For 20 nM EGFR, they were $(65.2 \pm 0.2)\%$ and $(85.1 \pm 0.2)\%$, respectively ($N = 924$). For 40 nM EGFR, they were $(65.3 \pm 0.2)\%$ and $(85.5 \pm 0.2)\%$, respectively ($N = 1711$). **c** Histograms (in blue) of the EGFR-released durations (τ_{on}) at various EGFR concentrations, [EGFR]. τ_{on} (mean \pm s.e.m.) were 0.78 ± 0.04 s (number of events: $N = 491$), 0.42 ± 0.03 s ($N = 843$), and

0.25 ± 0.02 s ($N = 1641$) at [EGFR] values of 10 nM, 20 nM and 40 nM, respectively. Fits are indicated in red. **d** Histograms (in green) of EGFR-captured durations (τ_{off}) at different [EGFR]. The cumulative fits are marked in black. The red and cyan curves indicate fits for short- and long-lived EGFR captures. For 10 nM EGFR, they (mean \pm s.e.m.) were 0.072 ± 0.011 s and 1.2 ± 0.1 s, respectively (number of events: $N = 441$). For 20 nM EGFR, they were 0.069 ± 0.007 s and 0.96 ± 0.09 s, respectively ($N = 806$). For 40 nM EGFR, they were 0.066 ± 0.006 s and 0.81 ± 0.11 s, respectively ($N = 1598$). **e** Dependence of the event frequencies in the form of $1/\tau_{on-i}$ on [EGFR], where τ_{on-1} (in red) and τ_{on-2} (in green) are the released durations between the short and long captures, respectively. **f** Dependence of $1/\tau_{off-i}$ on [EGFR], where $i = 1$ (in red) and $i = 2$ (in green) are subscripts corresponding to the short- and long-lived EGFR captures, respectively. Data points in panels **e** and **f** represent mean \pm s.d. obtained from $n = 3$ different experiments. Source data are provided as a Source Data file.

and 10). This outcome suggests two distinct mechanisms of binding of EGFR to Adnectin1, which correlate with the extent of the normalized current amplitude of EGFR-captured events and their duration. The event frequencies of short-lived and long-lived EGFR-captured events, in the form of $1/\tau_{\text{on-1}}$ and $1/\tau_{\text{on-2}}$, respectively, were linearly dependent on the EGFR concentration, [EGFR] (Fig. 4e). Here, $\tau_{\text{on-1}}$ and $\tau_{\text{on-2}}$ are the release (e.g., interevent) durations corresponding to the short-lived and long-lived current blockades, respectively (Supplementary Table 12). The slopes of the linear fits of $1/\tau_{\text{on-}i}$ ($i = 1, 2$) versus [EGFR] are the association rate constants, $k_{\text{on-}i}$, of Adnectin1-EGFR interactions because $k_{\text{on-}i} = 1/(\tau_{\text{on-}i}[\text{EGFR}])$. Again, the dissociation constants of the short-lived ($k_{\text{off-1}}$) and long-lived ($k_{\text{off-2}}$) current blockades were independent of [EGFR] (Fig. 4f; Supplementary Table 14). We interpret that these blockades are produced by specific bindings of EGFR to Adnectin1. We obtained the association rate constants, $k_{\text{on-1}}$ and $k_{\text{on-2}}$ (mean \pm s.e.m.), of $(6.62 \pm 0.21) \times 10^7 \text{ M}^{-1} \text{ s}^{-1}$ and $(2.89 \pm 0.10) \times 10^7 \text{ M}^{-1} \text{ s}^{-1}$, respectively. The dissociation rate constants, $k_{\text{off-1}}$ and $k_{\text{off-2}}$ (mean \pm s.e.m.), were $12.0 \pm 0.4 \text{ s}^{-1}$ and $1.01 \pm 0.01 \text{ s}^{-1}$, respectively (Supplementary Table 15). These values yield the equilibrium dissociation constants of the short-lived and long-lived current blockades, $K_{\text{D-1}}$ and $K_{\text{D-2}}$ (mean \pm s.e.m.), of $181 \pm 8 \text{ nM}$ and $34 \pm 2 \text{ nM}$, respectively.

The EGFR structure in the EGFR/EGF complex (“1NQL”)⁵⁵ is similar to that of EGFR in the EGFR-Adnectin1 complex (“3QWQ”)⁴⁷. It is believed to be an inactive receptor form (Supplementary Fig. 17a, b)^{47,55}. Adnectin1 and EGF bind to the EGFR domain D-I with a highly overlapping binding surface (Supplementary Fig. 17c, d). EGFR is a remarkably adaptable molecule with relatively rigid domains D-I and D-III^{55,56}. In contrast, the domains D-II and D-IV can adopt multiple conformations that place domain D-III differently in relation to domain D-I. We speculate that such distinct conformers of a flexible EGFR may likely be responsible for the bimodal protein recognition of EGFR by Adnectin1. The extended time bandwidth of our measurements facilitated the detection and quantification of conformational binding substates of the EGFR-Adnectin1 complex that are hidden in ensemble or low-resolution single-molecule measurements⁵⁷. Earlier studies using the resistive-pulse technique have also reported multimodal conformational transitions in the case of the dihydrofolate reductase (DHFR) enzyme²¹.

Are there interconversion transitions between the capture substates?

Next, we asked whether these reversible current transitions may also involve transitions between the two EGFR-captured substates. Hence, a related question is whether a kinetic model including interconversion transitions between these EGFR-captured substates would more accurately reflect experimentally determined rate constants. An interconversion-dependent kinetic model was developed, encompassing two supplementary rate constants between EGFR-captured substates, k_{12} and k_{21} (Supplementary Table 16 and Fig. 18). At a confidence level of $C > 0.95$, we found that fits to an interconversion-dependent kinetic model were not statistically superior over those corresponding to an interconversion-independent kinetic model, as indicated by the LLR test. Finally, to examine the reactivity crosscheck of our sensors, we recorded electrical traces of Adnectin1-tFhuA in the presence of either hSUMO1 (Supplementary Fig. 19) or WDR5 (Supplementary Fig. 20). In both cases, very short-lived and low-amplitude current blockades were noted. These blockades resemble those typically found in the case of nonspecific interactions of folded proteins with the *cis* opening of tFhuA (Supplementary Figs. 9, 12, and 16). This finding proves that the Adnectin1-tFhuA sensor is highly specific to EGFR.

Single-molecule detection of a protein biomarker in a biofluid

We challenged this sensor in the presence of 5% (v/v) fetal bovine serum (FBS) to examine the stability of this system in a harsh environment and the ability to distinguish analyte-captured events from

other nonspecific transitions of the solution constituents. Under physiological conditions, the serum threshold for the soluble EGFR ectodomain level is 45 ng/ml ($\sim 112 \text{ nM}$)⁵⁸. The tumor state can be evaluated at EGFR levels significantly exceeding this threshold⁵⁴. Fig. 5a, b show a representative signature of Adnectin1-tFhuA without and with 20 nM EGFR, respectively. However, the addition of 5% (v/v) FBS decorated the standard signature of EGFR-captured events with brief current spikes in the low-millisecond range (Fig. 5c). The power spectral density (PSD) of current fluctuations revealed a transition from white noise in the absence of FBS to $1/f$ flicker noise in the presence of FBS (Fig. 5d). This outcome suggests low-frequency equilibrium fluctuations in the local mobility and density of charges at the nanopore tip in the presence of FBS⁵⁹. The brief FBS-induced current fluctuations had a lower current amplitude around the open O_{on} substate (Fig. 5e–g), indicating that these may result from trafficking moieties of serum constituents at the *cis* opening of Adnectin1-tFhuA (Fig. 1). An extensive statistical analysis of the current blockades corresponding to the $O_{\text{off-1}}$ and $O_{\text{off-2}}$ levels confirmed the presence of two EGFR-captured event types in the presence of FBS (Fig. 5h–k; Supplementary Tables 17 and 18). No statistically significant impact of FBS was noted on the $k_{\text{off-1}}$ and $k_{\text{off-2}}$, but small changes, within the same order of magnitude, on the $k_{\text{on-1}}$ and $k_{\text{on-2}}$ (Supplementary Table 19). These changes may result from the interference of serum constituents with the binding interfaces of EGFR and Adnectin1. The mean duration of long-lived EGFR-induced current blockades was $\tau_{\text{off-2}} = 0.93 \pm 0.14 \text{ s}$, much longer than the brief millisecond-timescale FBS-induced closures. Under these conditions, we determined a corresponding $\tau_{\text{on-2}}$ of $1.7 \pm 0.5 \text{ s}$. Using a $k_{\text{on-2}}$ of $(2.9 \pm 0.2) \times 10^7 \text{ M}^{-1} \text{ s}^{-1}$ in the absence of FBS, we can evaluate the EGFR concentration in the serum sample, $[\text{EGFR}]_{\text{FBS}}$, using the equation $[\text{EGFR}]_{\text{FBS}} = 1/(\tau_{\text{on-2}} k_{\text{on-2}})$. Employing these values, we determined an $[\text{EGFR}]_{\text{FBS}}$ of $22.2 \pm 5.9 \text{ nM}$ in the FBS-containing sample, which is near the actual concentration of 20 nM.

Distinct outcomes with the monobody-based sensors

In this study, we provide a detailed signature analysis of single-molecule protein detection of three analytes using three nanopore sensors that share a modular architecture but differ by their binding surface (Supplementary Figs. 21–23). Fortuitously, all monobodies partly block the ionic flow through tFhuA, allowing direct electrical detection of analyte bindings without needing any peptide tag^{32,33}. As the protein analytes and their complexes with the specific monobodies drastically vary in size, charge, and structural complexity, distinct current blockades are noted in each case (Fig. 6; Supplementary Tables 1 and 20). For example, WDR5 interacts with a distal FG loop of Mb4 and away from the tFhuA pore opening (Supplementary Fig. S22), suggesting a modest current blockade made by the WDR5-Mb4 complex. In accord with this expectation, we note low-amplitude current blockades produced by WDR5-captured events (Fig. 3a; Supplementary Table 20). In contrast, the conformational complexity and structural properties of the hSUMO1-FN3SUMO and EGFR-Adnectin1 complexes at the tip of tFhuA indicated a potentially large current blockade, as also found by electrical recordings (Fig. 2a and Fig. 4a). In addition, we probed distinct single-molecule kinetic signatures of each analyte without the steric restrictions of the nanopore confinement (Supplementary Fig. 24). These unique characteristics of protein detections using externally engineered complex binding interfaces culminated with the discovery of bimodal protein recognition of EGFR.

Validation of the monobody-based sensors

Next, we examined the binding affinity of detergent-refolded sensors with their cognate analytes using steady-state fluorescence polarization (FP) anisotropy. If the labeled protein analyte interacts with the corresponding monobody-containing sensor, its tumbling rate (e.g., the coefficient of rotational diffusion) decreases, increasing the FP anisotropy. In accordance with our prediction, the FP anisotropy

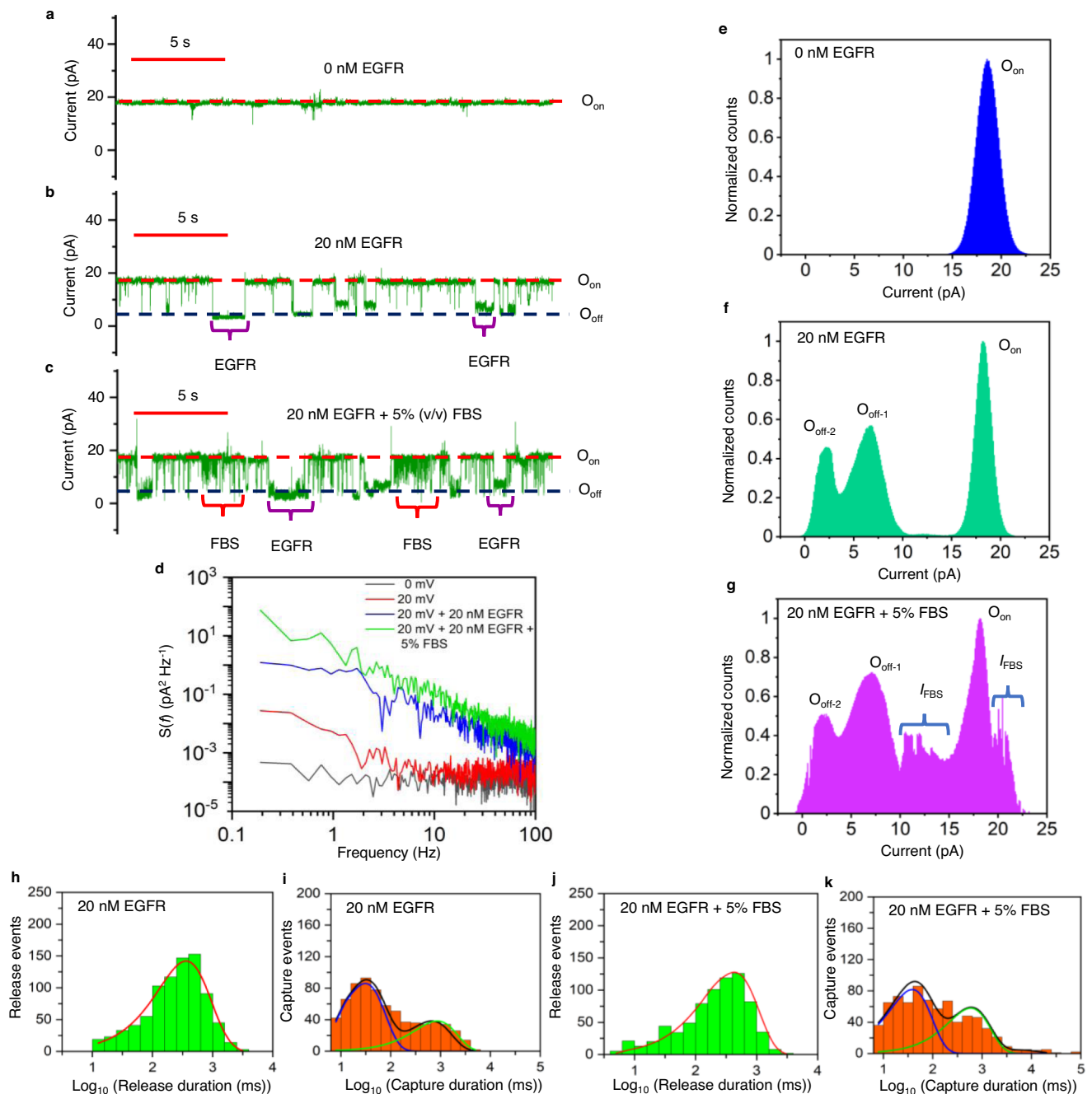


Fig. 5 | Single-molecule detection and quantification of EGFR in a biofluid. **a** An electrical trace of Adnectin1-tFhuA. **b** The trace in **a** with 20 nM EGFR. **c** The trace in **b** with 5% (v/v) FBS. The transmembrane potential was +20 mV. This subset of single-channel electrical signatures (in green) was replicated in $n = 3$ independent experiments. O_{on} and O_{off} are the EGFR-released and EGFR-captured substates, respectively. Traces were filtered at 0.5 kHz. **d** Power spectral density of current noise ($S(f)$) of traces from **a–c**. Each spectrum represents an average of three independent traces. **e** A current histogram (in blue) of the O_{on} substate of Adnectin1-tFhuA. The amplitude (mean \pm s.e.m.) of the O_{on} substate was 18.2 ± 0.1 pA. **f** A current histogram (in green) of the O_{on} and O_{off} substates of Adnectin1-tFhuA at 20 nM EGFR. The currents (mean \pm s.e.m.) of the O_{off-1} and O_{off-2} substates were 6.4 ± 0.1 pA and 2.3 ± 0.1 pA, respectively. **g** A current histogram (in magenta) of the O_{on} and O_{off} substates of Adnectin1-tFhuA at 20 nM EGFR and in the presence of 5% fetal bovine serum (FBS). This plot reveals the residual signal

produced by the FBS constituents (I_{FBS}). **h** A histogram (in green) of the EGFR-released durations (τ_{on}) at 20 nM EGFR. τ_{on} (mean \pm s.e.m.) was 0.40 ± 0.03 s (number of events: $N = 844$). **i** A histogram (in orange) of the EGFR-captured durations (τ_{off}) at 20 nM EGFR. τ_{off-1} and τ_{off-2} (mean \pm s.e.m.) were 0.044 ± 0.015 s and 0.982 ± 0.049 s, respectively (number of events: $N = 734$). **j** A histogram (in green) of the EGFR-released durations (τ_{on}) at 20 nM EGFR and in the presence of 5% FBS. τ_{on} (mean \pm s.e.m.) was 0.607 ± 0.051 s ($N = 738$). **k** A histogram (in orange) of the EGFR-captured durations (τ_{off}) at 20 nM EGFR and in the presence of 5% FBS. τ_{off-1} and τ_{off-2} (mean \pm s.e.m.) were 0.036 ± 0.013 s and 0.81 ± 0.08 s, respectively ($N = 694$). In panels **h** and **j**, the fits are indicated in red. In panels **i** and **k**, the cumulative fits are marked in black. The blue and green curves indicate fits for the short- and long-lived EGFR captures, respectively. Source data are provided as a Source Data file.

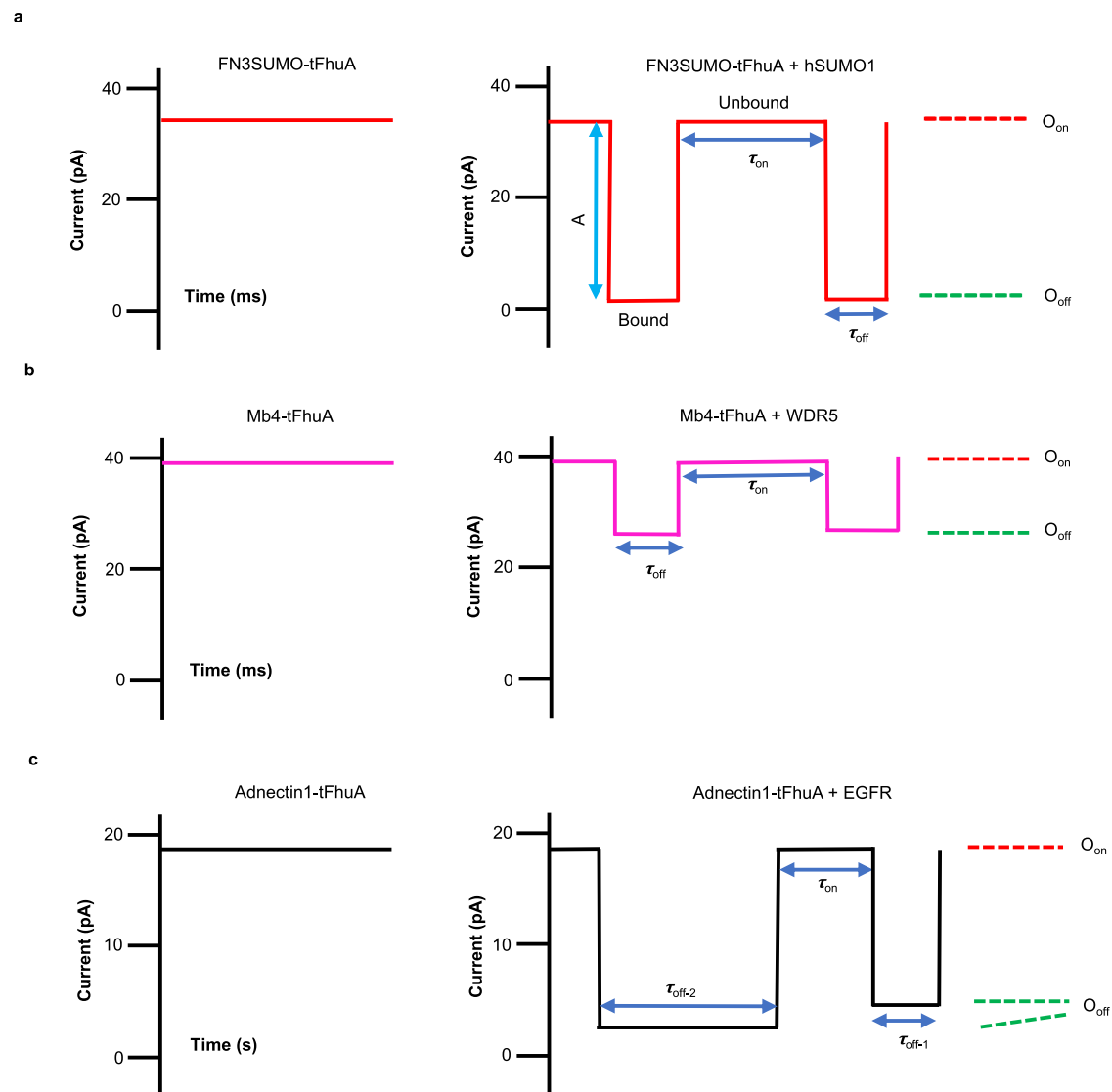


Fig. 6 | Schematic of the relative current blockades acquired with FN3SUMO-tFhuA, Mb4-tFhuA, and Adnectin1-tFhuA. O_{on} and O_{off} are the protein analyte-released and protein analyte-captured substates, respectively. τ_{on} and τ_{off} are the durations of the protein analyte-released and protein analyte-captured events, respectively. **a** Graphic representation of stochastic sensing of hSUMO1 using an FN3SUMO-tFhuA sensor, which maintains an open-state current (left panel). When hSUMO1 is added to the *cis* side, the analyte produces large-amplitude (A), medium-lived current transitions between two current substates (right panel).

b Mb4-tFhuA sensor maintains an open-state current (left panel). When added to the *cis* side, WDR5 produces low-amplitude, medium-lived current transitions (right panel). **c** Adnectin1-tFhuA protein maintains an open-state current (left panel). When added to the *cis* side, EGFR produces short- and long-lived current transitions (right panel). Note that there are two EGFR-captured durations, τ_{off-1} and τ_{off-2} . The short-lived transitions feature a relatively lower current amplitude than the long-lived transitions.

substantially increased at elevated sensor concentrations (Supplementary Fig. 25). Conversely, tFhuA at increased concentrations did not alter the FP anisotropy, confirming no interaction between labeled proteins and tFhuA-containing detergent micelles. The calculated K_D values of hSUMO1 and WDR5 with their respective nanopore sensors were 186 ± 16 nM and 455 ± 59 nM, respectively. These data agree well with those derived from single-channel electrical recordings (Supplementary Tables 5 and 8). EGFR is unsuitable for this assay because of its large molecular weight, producing a longer tumbling rate than the fluorescence lifetime of most fluorophores. However, the K_D determined for the long-lived EGFR-captured events using Adnectin1-tFhuA sensor is in accord with a previously reported study (Supplementary Table 15)⁴⁷. It should be mentioned that restraining one binding partner to a surface can decrease the affinity by one order of magnitude⁵³. Hence, this explains a significantly weaker binding interaction with the monobody-containing nanopore

sensor immobilized on a lipid bilayer than that value measured in solution by steady-state FP spectroscopy.

Advantages of these nanopore sensors and their implications in nanobiotechnology

In this study, we engineered nanopore sensors made of a single-polypeptide unit that feature a selective protein binder adaptable with atomic precision. The monomeric nature of these sensors circumvents the necessity of tedious purification steps of the assembly reaction, otherwise required for multimeric nanopores. The overall architecture of the sensors can be maintained while changing the interaction interface of the antibody-mimetic binder. This way, such an approach substantially extends the applications of these sensing elements for numerous protein biomarkers. This critical benefit is facilitated by the genetically encoded nature of these sensors so that they can create combinatorial libraries of tethered binders. For instance, the loops of

monobodies are analogous to the complementarity-determining regions (CDRs) of antibodies. One significant advantage of monobodies is their ability to interact with challenging binding surfaces that are not generally exposed to the CDRs of antibodies³⁴. In addition, there is no fundamental limitation in replacing the monobody with another synthetic protein binder^{60–62}. Furthermore, the main benefits of using antibody-mimetic proteins include strong binding affinities with different epitopes, straightforward expression and purification procedures, and high thermodynamic stability³⁴.

A shortcoming of most detection techniques is the nonspecific binding of assay reagents and immobilization surfaces. This problem determines an increase in background noise, interfering with the signal produced by protein binding events. Enzyme-linked immunosorbent assay (ELISA) is the most common protein detection technique in basic research and molecular diagnostics. This technology is also formulated as single-molecule arrays that rely on bead-based, femtoliter-size wells^{1,63}. The significant advantage of ELISA is its high detection performance in the pg/ml range. However, its dynamic range is limited to four orders of magnitude.

Notably, our method has the potential to detect and characterize functionally distinct subpopulations of specific binding events in a challenging biofluid. This is a significant advantage that existing technologies in the ensemble lack, such as BLI, surface plasmon resonance (SPR), isothermal titration calorimetry (ITC), and ELISA. BLI, SPR, and nanopore sensors are amenable to parallel recording technologies, but ITC cannot be used at high throughput and necessitates large amounts of proteins. Our modular sensors rely on a single-molecule detector that enables direct measurements of time-resolved protein binding events at adjustable protein concentrations. These nanopore sensors feature high-affinity protein binders that do not cross-react. In addition, our sensors may be further developed for concurrently identifying two or more target proteins of varying binding affinity and specificity because each protein would provide a specific signature. These competitive binding interactions against the same protein binder cannot be analyzed using bulk-phase techniques. We probe the complexity and heterogeneity of protein recognition events without requiring any additional exogenous tag or peptide tail. In addition, this sensor formulation includes a system that precludes the occurrence of nonspecific binding events or protein inactivation at the liquid-surface interface, as in the case of surface immobilization-based sensors. Our proposed approach shows prospects for discovering rare and short-lived binding events, which are unlikely to be detectable by prevailing technologies. In extreme conditions of unusually high k_{on} , such as those in the range of 10^7 – 10^9 M⁻¹s⁻¹¹²⁹, we show that our method can be utilized to measure such values (e.g., for hSUMO1 and WDR5). In nanopore-based sensing, the k_{off} can be recorded up to a value of $\sim 10^5$ s⁻¹. In contrast, SPR and BLI have challenges detecting this value greater than 1 s⁻¹. Nanopore sensors with external protein recognition elements may detect protein levels at least two orders of magnitude below the K_D in solution³¹. If these nanopores encompass very high-affinity binders, such as affibodies, with a K_D in a low-picomolar range⁶⁴, then their detection sensitivity may attain pg/ml levels. Hence, these sensors can operate at clinically relevant concentration ranges of proteins and with an extended time bandwidth of at least eight orders of magnitude. In this process, the analyte-induced events are unambiguously distinguished from other nonspecific current blockades of biofluid constituents. With further developments, these sensors can be integrated with high-throughput technologies for biomarker profiling in biomedical diagnostics.

Methods

Computational grafting of monobodies onto tFhuA

For the structural prediction of nanopore sensors, the amino acid sequence of each monobody (FN3, FN3SUMO, Mb4, and Adnectin1) was inserted at the N-terminus of tFhuA via a (GGG)₂ peptide tether^{31,33}.

3D structural models of the nanopore sensors were generated in silico using AlphaFold2^{48,49}. All parameters were kept the same for all nanopore sensors. The predicted structures of sensors were confirmed by comparisons with individual structures of FhuA and monobodies.

Synthetic gene construction

Three derivatives of wild-type fibronectin type-III (FN3) were used to develop these sensors. The cDNA sequences of these *fn3* genes, namely *fn3sumo*, *mb4*, and *adnectin1*, were fused to the 5' end of the *tfhua* gene via a (GGG)₂-encoding linker by a restriction-free cloning method⁶⁵. The cDNA sequences of Mb4 and Adnectin1 were synthesized by Eurofins Genomics (Louisville, KY) and Integrated DNA Technologies (IDT, Coralville, Iowa), respectively. The construction of the *fn3sumo* gene was made based on ySMB9⁴⁵. The cDNA sequence of all three fibronectin derivatives was first amplified using Q5 high-fidelity DNA polymerase (New England Biolabs, Ipswich, MA) from their respective template DNA. PCR products were separated on 1% agarose gel and purified using a Gel extraction kit (Promega, CA). Sequences of forward and reverse primers are listed in Supplementary Table 21. Amplified products of *fn3sumo* and *mb4* genes were then fused to the 5' end of *tfhua* cloned in pPR-IBA1 plasmid (IBA, Goettingen, Germany). *adnectin1* was joined at the 5' end of *tfhua* in pET28a (EMD Millipore, Burlington, MA). The pET28-tFhuA plasmid was constructed by inserting the gene between *Bam*HI and *Xho*I restriction sites after amplification with forward and reverse primers of tFhuA (Supplementary Table 21). All the gene sequences were verified by sequencing (MCLab, San Francisco, CA). The pET11a-hSUMO1 was kindly provided by Fauke Mechior (Addgene plasmid #53138).

Protein expression and purification

For the expression of FN3SUMO-tFhuA, Mb4-tFhuA, and Adnectin1-tFhuA, the plasmids mentioned above were transformed into *E. coli* BL21(DE3) cells. These monobody-containing protein nanopores were purified from inclusion bodies. Cells containing these overexpressed protein nanopores were lysed using a microfluidizer (Model 110L; Microfluidics, Newton, MA). Then, cells underwent a series of washes in 1% (v/v) Triton. Cell lysates were centrifuged at $108,500 \times g$ for 30 min at 4 °C. This way, the insoluble pellet and supernatant were separated. The insoluble pellet was solubilized in 8 M urea. This was followed by a first purification step achieved through anion-exchange chromatography (Q12-Sepharose; Bio-Rad, Hercules, CA). Then, protein samples were purified further using size-exclusion chromatography (HiLoad 16/600 Superdex-75; GE Healthcare Life Sciences, Pittsburg, PA). The protein purity was validated by sodium dodecyl-sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) analysis (Supplementary Fig. 1). No noteworthy differences were detected among the productions of the three sensors. In the case of hSUMO1, BL21(DE3) cells were transformed with pET11a-hSUMO1 and grown in Luria-Bertani (LB) medium at 37 °C until OD₆₀₀ attained a value of ~ 0.5 . Then, the temperature was changed to 20 °C. Expression was initiated by inducing the cells with 250 μM IPTG. After induction, the cells were cultured for ~ 18 h at 20 °C. Cells were centrifuged at $3700 \times g$ for 30 min at 4 °C, then resuspended in 50 mM Tris-HCl, 50 mM NaCl, and pH 8.0. The lysozyme was added to the suspended cells and incubated on ice for 15 min. Cell lysis was accomplished using sonication (30 s on, 60 s off $\times 4$ times). The cell lysate was centrifuged at $108,500 \times g$ for 30 min at 4 °C. Then, the supernatant was filtered using a 0.22 μm filter. The supernatant was loaded onto a Q-Sepharose column (Cytiva, Marlborough, MA), which was washed with 50 mM Tris-HCl, 50 mM NaCl, pH 8.0, and eluted with 50 mM Tris-HCl, 1 M NaCl, pH 8.0 in a gradient manner. The desired fractions were collected, dialyzed, and concentrated. Furthermore, the protein sample was loaded on an S75 gel-filtration column (GE Healthcare, Chicago, IL). Pure fractions were collected and dialyzed against 20 mM Tris-HCl, 150 mM NaCl, pH 8.0, and 0.5 mM TCEP overnight at 4 °C.

WDR5 was expressed using ROSETTA cells (New England Biolabs, Ipswich, MA). Cells were induced utilizing 100 μ M IPTG. The cell lysis was conducted using a microfluidizer (Model 110L; Microfluidics, Newton, MA). The lysates were centrifuged at 3700 $\times g$ for 30 min at 4 °C. The WDR5-containing supernatant underwent an initial purification process via a metal-affinity chromatography (5 ml, Bio-Scale Mini Profinity IMAC cartridge; Bio-Rad, Hercules, CA). Then, a Tobacco Etch Virus (TEV) protease (New England Biolabs) enzymatic assay was used to remove the hexahistidine tag, which was followed by the addition of the benzonase nuclease (Sigma-Aldrich, St. Louis, MO) to digest DNA contaminants. Finally, the sample was again passed through a metal-affinity column to remove protein aggregates. Finally, a 10 kDa-molecular weight concentrator (Millipore Sigma, St. Louis, MO) was utilized to prep the final protein samples.

For the purification of the ectodomain of epidermal growth factor receptor (EGFR), Expi293F cells (Thermo Fischer Scientific; Cat# A14527) were seeded at 10⁶ cells/ml density in 1 l of Dynamis growth medium (Gibco) 24 h before the transfection and supplemented with Tryptone/Glucose. For the sake of simplicity, we name this EGFR throughout this article. The culture was transfected with 2 μ g/ml of the pCMV_EGFR plasmid containing the signal peptide with 3.75 \times poly-ethylenimine (PEI). Transfected cells were cultured for five days, and the protein was allowed to excrete from the cells. Five days post-transfection, the culture was pelleted, and the supernatant was filtered. The sample was loaded onto an immobilized metal-affinity column (1 ml, HiStrap HP column, GE Healthcare), which was washed with 50 mM sodium phosphate (NaPi) (pH 8.0), 300 mM NaCl, 20 mM imidazole. The protein was eluted using 50 mM NaPi (pH 8.0), 300 mM NaCl, and 500 mM imidazole. Peak fractions were collected and confirmed by SDS-PAGE (Supplementary Fig. 26). Finally, the protein sample was concentrated and exchanged with phosphate buffer saline (PBS, pH 7.5) using a PD10 column (GE Healthcare) and stored at -80 °C. The purity of all protein analytes was tested by SDS-PAGE analysis.

Protein refolding

The purified FN3SUMO-tFhuA, Mb4-tFhuA, and Adnectin1-tFhuA were adjusted to a final concentration of -10 μ M. Next, n-dodecyl- β -D-maltopyranoside (DDM) was added to denatured samples to a final concentration of 1% (w/v). The protein samples were immediately dialyzed against the buffer containing 200 mM KCl, 20 mM Tris-HCl, pH 8, at 4 °C for 96 h. The dialysis solution was replaced at 24-h intervals. These refolded protein samples were centrifuged to eliminate any protein precipitations, and the supernatant was used as the running sample for single-channel electrical recordings. Protein concentrations were determined by their molar absorptivity at a wavelength of 280 nm.

Single-channel electrical recordings

Electrical detection of protein ligands at single-molecule precision was conducted using planar lipid bilayers⁶⁶. The two halves of the chamber were divided by a 25 μ m-thick Teflon septum (Goodfellow Corporation, Malvern, PA). A planar lipid bilayer was made of 1,2-diphytanoyl-*sn*-glycero-phosphatidylcholine (Avanti Polar Lipids, Alabaster, AL) across an -100 μ m-diameter aperture of the Teflon septum. For all experiments, the buffer solution contained 300 mM KCl, 10 mM Tris-HCl, and pH 8.0. In addition, this buffer included 0, 0.5, and 1 mM TCEP in experiments with EGFR, hSUMO1, and WDR5, respectively. The nanopore protein samples (final concentration, 0.5–1.5 ng/ μ l) and analytes were added to the *cis* compartment, which was grounded. Single-channel electrical currents were acquired using an Axopatch 200B patch-clamp amplifier (Axon Instruments, Foster City, CA) controlled by Clampex 10.7 (Axon Instruments). The applied transmembrane potential was +40 mV, unless otherwise stated. The electrical signal was sampled at 50 kHz using a low-noise acquisition system (Model Digidata 1440 A; Axon Instruments). A low-pass Bessel filter (Model 900; Frequency

Devices, Ottawa, IL) was further employed for signal filtering at 10 kHz. For the data processing and analysis, the electrical traces were digitally filtered with a low-pass 8-pole Bessel filter at 3 kHz, unless otherwise stated. All single-channel electrical recordings were acquired at a temperature of 24 \pm 1 °C.

EGFR detection in a heterogeneous solution

For the detection and quantification of EGFR in heterogeneous solutions, fetal bovine serum (FBS, Gibco™, Thermo Fisher Scientific, Pittsburgh, PA) was used. FBS was sterilized through a syringe filter before being stored at -80 °C. For single-channel recording, an aliquot was defrosted on ice and kept at room temperature before adding to the chamber. Single-channel electrical traces were recorded in the presence of FBS at a final concentration of 5% (v/v). At this FBS concentration, the lipid bilayers were stable, and no clogging of nanopores was noted for long recording periods³². Under these conditions, the nontarget FBS protein concentration for analyzing single-molecule EGFR detection was 1.5–2.5 mg/ml (provided by Gibco™). These traces were filtered with a low-pass 8-pole Bessel filter at 500 Hz.

Biolayer interferometry (BLI) assay using immobilized proteomicelles

These experiments were conducted using an Octet Red384 instrument (FortéBio, Fremont, CA) at 24 °C⁶⁷. For BLI experiments, a site-specific insertion of cysteine at position 287 was achieved in the long L4 loop of Mb4-tFhuA by site-directed mutagenesis (Q5 mutagenesis kit, New England Biolabs). This cysteine-containing Mb4-tFhuA was expressed and purified as described above, except for the presence of a reducing agent. Cys287 was biotinylated using maleimide chemistry. A flexible (PEG)₁₁ linker was used between the biotin and maleimide groups. The BLI running buffer contained 300 mM KCl, 20 mM Tris-HCl, 1 mM TCEP, 1% DDM, 1 mg/ml bovine serum albumin (BSA), pH 8.0. It was used to soak streptavidin (SA) sensors for 30 min. The 50 nM Mb4-tFhuA_Cys287-(PEG)₁₁-Biotinyl was loaded onto the sensors for 2.5 min via biotin-streptavidin chemistry. By dipping the sensors in a protein-free solution for 6 min, the unattached Mb4-tFhuA_Cys287 was washed away. The association process was examined using various concentrations of WDR5, ranging from 1.5 μ M to 6 μ M. The BLI sensors were dipped in a WDR5-free running buffer to inspect the dissociation phase. For all WDR5 concentrations, the Mb4-tFhuA_Cys287-free BLI sensors were run in parallel as controls. The baseline and drift in the sensorgrams were subtracted using these controls. FortéBio Octet HT acquisition 11 and FortéBio Octet HT kinetic analysis 11 (FortéBio) were used for data acquisition and analysis, respectively.

Steady-state fluorescence polarization (FP) measurements

hSUMO1 and WDR5 were labeled with fluorescein and rhodamine, respectively, at pH 9.0 by primary amine chemistry. These labeled proteins were added to the well at a final concentration of 50 nM. Steady-state fluorescence polarization (FP) anisotropy assays were conducted in triplicate with an 18-point serial dilution of FN3SUMO-tFhuA, Mb4-tFhuA, or unmodified tFhuA, against a fixed concentration of labeled proteins on black 96-well plates. All steady-state FP measurements were recorded using a SpectraMax i3x plate reader (Molecular Devices, San Jose, CA) controlled by SoftMax Pro 6.4 (Molecular Devices). All measurements were conducted at 0 min and after a one-hour incubation at room temperature in the dark. The resulting dose-response data were averaged and fitted using logistic regression to obtain each interaction's dissociation constant (K_D).

Statistics and reproducibility

pClamp 10.7 (Axon Instruments) was used for the data acquisition and analysis. Capture and release events were collected using single-channel event searches by ClampFit 10.7 (Axon Instruments), and

figures were prepared by Origin 9.7 (OriginLab, Northampton, MA). The probability distribution function (PDF) was generated using a kinetic rate matrix, and the kinetic rate constants were determined by fitting the data using the maximum likelihood method⁵¹. To evaluate the results of multiple models and select the number of statistically significant peaks that are best matched to the data, a logarithm likelihood ratio (LLR) test was performed⁵². At a confidence number of $C = 0.95$, a single-exponential fit was the best model for the release and capture durations of hSUMO1 and WDR5. For EGFR, a two-exponential fit was the best model for the capture durations. Single-channel traces were recorded for long periods, so a single statistical analysis encompassed several hundred events. This way, satisfactory reproducibility was achieved from one experiment to the next, as indicated by values of standard deviations. These statistical analyses were not subjected to randomizing and blinding.

Molecular graphics

All cartoons showing molecular graphics were prepared using PyMOL (Version 2.4.0; Schrödinger, LLC) and Chimera X (Version 1.4; The University of California at San Francisco). Entries “3RZW”, “6BYN”, “1NQL”, “3QWQ”, “1BY3”, and “1FNF” from Protein Data Bank were used in this article for visualizations and molecular graphics of hSUMO1-FN3SUMO1, WDR5-Mb4, EGFR ectodomain, EGFR-Adnectin1, tFhuA, and FN3, respectively.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

In addition to the Supplementary Information file, data supporting the findings of this article have been deposited in the Zenodo database at <https://doi.org/10.5281/zenodo.7504671>. The source data underlying Figs. 2b–g, 3b–g, 4b–f, 5d–k, and Supplementary Figs. 6–7, 14b, 24–25 and Supplementary Tables 2–20 are provided in the Source data file. Entries “3RZW”, “6BYN”, “1NQL”, “3QWQ”, “1BY3”, and “1FNF” from Protein Data Bank were used in this article for visualizations and molecular graphics. Source data are provided with this paper.

Code availability

All custom codes and mathematical algorithms were based on MATLAB R2021a. They are available at Zenodo [<https://doi.org/10.5281/zenodo.7504671>].

References

1. Cohen, L. & Walt, D. R. Highly sensitive and multiplexed protein measurements. *Chem. Rev.* **119**, 293–321 (2019).
2. Kustatscher, G. et al. Understudied proteins: opportunities and challenges for functional proteomics. *Nat. Methods* **19**, 774–779 (2022).
3. Chatterjee, T. et al. Direct kinetic fingerprinting and digital counting of single protein molecules. *Proc. Natl Acad. Sci. USA* **117**, 22815–22822 (2020).
4. Mandal, S. et al. Direct kinetic fingerprinting for high-accuracy single-molecule counting of diverse disease biomarkers. *Acc. Chem. Res.* **54**, 388–402 (2021).
5. Wang, Y. et al. Single-molecule calorimeter and free energy landscape. *Proc. Natl. Acad. Sci. USA* **118**, e2104598118 (2021).
6. Ayub, M. & Bayley, H. Engineered transmembrane pores. *Curr. Opin. Chem. Biol.* **34**, 117–126 (2016).
7. Ying, Y. L., Yu, R. J., Hu, Y. X., Gao, R. & Long, Y. T. Single antibody-antigen interactions monitored via transient ionic current recording using nanopore sensors. *Chem. Commun. (Camb.)* **53**, 8620–8623 (2017).
8. Schmid, S. & Dekker, C. Nanopores: a versatile tool to study protein dynamics. *Essays Biochem.* **65**, 93–107 (2021).
9. Sackmann, B. & Neher, E. *Single-Channel Recording*, Edn. Second Edition (Kluwer Academic/Plenum Publishers, New York, 1995).
10. del Rio Martinez, J. M., Zaitseva, E., Petersen, S., Baaken, G. & Behrends, J. C. Automated formation of lipid membrane microarrays for ionic single-molecule sensing with protein nanopores. *Small* **11**, 119–125 (2015).
11. Cressiot, B., Bacri, L. & Pelta, J. The promise of nanopore technology: advances in the discrimination of protein sequences and chemical modifications. *Small Methods* **4** <https://doi.org/10.1002/smt.202000090> (2020).
12. Tanimoto, I. M. F. et al. Focus on using nanopore technology for societal health, environmental, and energy challenges. *Nano Res.* **15**, 9906–9920 (2022).
13. Ying, Y. L. et al. Nanopore-based technologies beyond DNA sequencing. *Nat. Nanotechnol.* **17**, 1136–1146 (2022).
14. Xing, Y., Dorey, A., Jayasinghe, L. & Howorka, S. Highly shape- and size-tunable membrane nanopores made with DNA. *Nat. Nanotechnol.* **17**, 708–713 (2022).
15. Piguet, F. et al. Identification of single amino acid differences in uniformly charged homopolymeric peptides with aerolysin nanopore. *Nat. Commun.* **9**, 966 (2018).
16. Ouldali, H. et al. Electrical recognition of the twenty proteinogenic amino acids using an aerolysin nanopore. *Nat. Biotechnol.* **38**, 176–181 (2020).
17. Yusko, E. C. et al. Real-time shape approximation and fingerprinting of single proteins using a nanopore. *Nat. Nanotechnol.* **12**, 360–367 (2017).
18. Rosen, C. B., Rodriguez-Larrea, D. & Bayley, H. Single-molecule site-specific detection of protein phosphorylation with a nanopore. *Nat. Biotechnol.* **32**, 179–181 (2014).
19. Versloot, R. C. A. et al. Quantification of protein glycosylation using nanopores. *Nano Lett.* **22**, 5357–5364 (2022).
20. Harrington, L., Alexander, L. T., Knapp, S. & Bayley, H. Single-molecule protein phosphorylation and dephosphorylation by nanopore enzymology. *ACS Nano* **13**, 633–641 (2019).
21. Galenkamp, N. S., Biesemans, A. & Maglia, G. Directional conformer exchange in dihydrofolate reductase revealed by single-molecule nanopore recordings. *Nat. Chem.* **12**, 481–488 (2020).
22. Rodriguez-Larrea, D. & Bayley, H. Multistep protein unfolding during nanopore translocation. *Nat. Nanotechnol.* **8**, 288–295 (2013).
23. Tripathi, P., Firouzbakht, A., Gruebele, M. & Wanunu, M. Threading single proteins through pores to compare their energy landscapes. *Proc. Natl Acad. Sci. USA* **119**, e2202779119 (2022).
24. Lucas, F. L. R., Versloot, R. C. A., Yakovlieva, L., Walvoort, M. T. C. & Maglia, G. Protein identification by nanopore peptide profiling. *Nat. Commun.* **12**, 5795 (2021).
25. Afshar Bakshloo, M. et al. Nanopore-based protein identification. *J. Am. Chem. Soc.* **144**, 2716–2725 (2022).
26. Alfaro, J. A. et al. The emerging landscape of single-molecule protein sequencing technologies. *Nat. Methods* **18**, 604–617 (2021).
27. Brinkerhoff, H., Kang, A.S.W., Liu, J., Aksimentiev, A. & Dekker, C. Multiple rereads of single proteins at single-amino acid resolution using nanopores. *Science* **374**, 1509–1513 (2021).
28. Rotem, D., Jayasinghe, L., Salichou, M. & Bayley, H. Protein Detection by Nanopores equipped with aptamers. *J. Am. Chem. Soc.* **134**, 2781–2787 (2012).
29. Harrington, L., Cheley, S., Alexander, L. T., Knapp, S. & Bayley, H. Stochastic detection of Pim protein kinases reveals electrostatically enhanced association of a peptide substrate. *Proc. Natl Acad. Sci. USA* **110**, E4417–E4426 (2013).

30. Fahie, M., Chisholm, C. & Chen, M. Resolved single-molecule detection of individual species within a mixture of anti-biotin antibodies using an engineered monomeric nanopore. *ACS Nano* **9**, 1089–1098 (2015).
31. Thakur, A. K. & Movileanu, L. Real-time measurement of protein-protein interactions at single-molecule resolution using a biological nanopore. *Nat. Biotechnol.* **37**, 96–101 (2019).
32. Thakur, A. K. & Movileanu, L. Single-molecule protein detection in a biofluid using a quantitative nanopore sensor. *ACS Sens.* **4**, 2320–2326 (2019).
33. Mayse, L. A. et al. Disentangling the recognition complexity of a protein hub using a nanopore. *Nat. Commun.* **13**, 978 (2022).
34. Koide, A., Wojcik, J., Gilbreth, R. N., Hoey, R. J. & Koide, S. Teaching an old scaffold new tricks: monobodies constructed using alternative surfaces of the FN3 scaffold. *J. Mol. Biol.* **415**, 393–405 (2012).
35. Kunz, P. et al. The structural basis of nanobody unfolding reversibility and thermoresistance. *Sci. Rep.* **8**, 7934 (2018).
36. Hantschel, O., Biancalana, M. & Koide, S. Monobodies as enabling tools for structural and mechanistic biology. *Curr. Opin. Struct. Biol.* **60**, 167–174 (2020).
37. Zhou, H. J. et al. SUMOylation of VEGFR2 regulates its intracellular trafficking and pathological angiogenesis. *Nat. Commun.* **9**, 3303 (2018).
38. Xie, M., Yu, J., Ge, S., Huang, J. & Fan, X. SUMOylation homeostasis in tumorigenesis. *Cancer Lett.* **469**, 301–309 (2020).
39. Ge, Z. et al. WDR5 high expression and its effect on tumorigenesis in leukemia. *Oncotarget* **7**, 37740–37754 (2016).
40. Chen, X. et al. Upregulated WDR5 promotes proliferation, self-renewal and chemoresistance in bladder cancer via mediating H3K4 trimethylation. *Sci. Rep.* **5**, 8293 (2015).
41. Cui, Z. et al. Effect of high WDR5 expression on the hepatocellular carcinoma prognosis. *Oncol. Lett.* **15**, 7864–7870 (2018).
42. Korpanty, G. J., Graham, D. M., Vincent, M. D. & Leigh, N. B. Biomarkers that currently affect clinical practice in lung cancer: EGFR, ALK, MET, ROS-1, and KRAS. *Front. Oncol.* **4**, 204 (2014).
43. Srivatsa, S. et al. EGFR in tumor-associated myeloid cells promotes development of colorectal cancer in mice and associates with outcomes of patients. *Gastroenterology* **153**, 178–190.e110 (2017).
44. Liu, X. et al. The importance of EGFR as a biomarker in molecular apocrine breast cancer. *Hum. Pathol.* **77**, 1–10 (2018).
45. Gilbreth, R. N. et al. Isoform-specific nanobody inhibitors of small ubiquitin-related modifiers engineered using structure-guided library design. *Proc. Natl Acad. Sci. USA* **108**, 7751–7756 (2011).
46. Gupta, A. et al. Facile target validation in an animal model with intracellularly expressed monobodies. *Nat. Chem. Biol.* **14**, 895–900 (2018).
47. Ramamurthy, V. et al. Structures of adnectin/protein complexes reveal an expanded binding footprint. *Structure* **20**, 259–269 (2012).
48. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
49. Tunyasuvunakool, K. et al. Highly accurate protein structure prediction for the human proteome. *Nature* **596**, 590–596 (2021).
50. Mohammad, M. M., Howard, K. R. & Movileanu, L. Redesign of a plugged beta-barrel membrane protein. *J. Biol. Chem.* **286**, 8000–8013 (2011).
51. Colquhoun, D., Hatton, C. J. & Hawkes, A. G. The quality of maximum likelihood estimates of ion channel rate constants. *J. Physiol.* **547**, 699–728 (2003).
52. Couoh-Cardel, S., Hsueh, Y. C., Wilkens, S. & Movileanu, L. Yeast V-ATPase proteolipid ring acts as a large-conductance transmembrane protein pore. *Sci. Rep.* **6**, 24774 (2016).
53. Imran, A. et al. Interplay of affinity and surface tethering in protein recognition. *J. Phys. Chem. Lett.* **13**, 4021–4028 (2022).
54. Banys-Paluchowski, M. et al. Evaluation of serum epidermal growth factor receptor (EGFR) in correlation to circulating tumor cells in patients with metastatic breast cancer. *Sci. Rep.* **7**, 17307 (2017).
55. Ferguson, K. M. et al. EGF activates its receptor by removing interactions that autoinhibit ectodomain dimerization. *Mol. Cell* **11**, 507–517 (2003).
56. Li, S. et al. Structural basis for inhibition of the epidermal growth factor receptor by cetuximab. *Cancer Cell* **7**, 301–311 (2005).
57. Teramura, Y. et al. Single-molecule analysis of epidermal growth factor binding on the surface of living cells. *EMBO J.* **25**, 4215–4222 (2006).
58. Sandri, M. T. et al. Serum EGFR and serum HER-2/neu are useful predictive and prognostic markers in metastatic breast cancer patients treated with metronomic chemotherapy. *Cancer* **110**, 509–517 (2007).
59. Sun, J., Thakur, A. K. & Movileanu, L. Protein ligand-induced amplification in the 1/f noise of a protein-selective nanopore. *Langmuir* **36**, 15247–15257 (2020).
60. Haussner, C., Lach, J. & Eichler, J. Synthetic antibody mimics for the inhibition of protein-ligand interactions. *Curr. Opin. Chem. Biol.* **40**, 72–77 (2017).
61. Yu, X., Yang, Y. P., Dikici, E., Deo, S. K. & Daunert, S. Beyond antibodies as binding partners: the role of antibody mimetics in bioanalysis. *Annu. Rev. Anal. Chem.* **10**, 293–320 (2017).
62. Meng, J. et al. Nanobody-based sensors reveal a high proportion of mGlu heterodimers in the brain. *Nat. Chem. Biol.* **18**, 894–903 (2022).
63. Cohen, L. et al. Single molecule protein detection with attomolar sensitivity using droplet digital enzyme-linked immunosorbent assay. *ACS Nano* **14**, 9491–9501 (2020).
64. Stahl, S. et al. Affibody molecules in biotechnological and medical applications. *Trends Biotechnol.* **35**, 691–712 (2017).
65. Unger, T., Jacobovitch, Y., Dantes, A., Bernheim, R. & Peleg, Y. Applications of the restriction free (RF) cloning procedure for molecular manipulations and protein expression. *J. Struct. Biol.* **172**, 34–44 (2010).
66. Larimi, M. G., Mayse, L. A. & Movileanu, L. Interactions of a polypeptide with a protein nanopore under crowding conditions. *ACS Nano* **13**, 4469–4477 (2019).
67. Imran, A. et al. Kinetics of the multitasking high-affinity Win binding site of WDR5 in restricted and unrestricted conditions. *Biochem. J.* **478**, 2145–2161 (2021).

Acknowledgements

We thank our colleagues in the Movileanu and Loh laboratories and at Ichor Life Sciences laboratories for their comments on the manuscript and their assistance during the very early stage of this project. We are grateful to Tom Duncan for his instrumental help in the initial phase of BLI experiments and Ali Imran for his generous assistance with the Q-matrix analysis. This work was supported by the National Institute of General Medical Sciences of the U.S. National Institutes of Health grants GM088403 (to L.M.), GM115762 (to S.N.L.), and GM148448 (to S.N.L.).

Author contributions

M.A., J.H.H., L.A.M., M.F.P., S.N.L., and L.M. designed research. M.A., J.H.H., M.F.P. performed research. M.A. and L.A.M. analyzed data. A.J.W., K.J.M., S.N.L., and L.M. provided research materials and funding. M.A., J.H.H., and L.M. wrote the paper.

Competing interests

M.A. and L.M. are named inventors on one pending provisional patent application (US 63/409,906) filed by Syracuse University on this work. The remaining authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-023-36944-9>.

Correspondence and requests for materials should be addressed to Liviu Movileanu.

Peer review information *Nature Communications* thanks Benjamin Cressiot, Vicente Aguilera and the other anonymous reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023