Single-molecule digital sizing of proteins in solution

Georg Krainer,[1,#] Raphael P.B. Jacquat,[1,#] Matthias M. Schneider,[1,#] Timothy J. Welsh,[1,#] Jieyuan Fan,[1,2,#] Quentin A. E. Peter,[1] Ewa A. Andrzejewska,[1] Greta Šneiderienė,[1] Magdalena A. Czekalska,[1] Hannes Ausserwoeger,[1] Lin Chai,[1] William E. Arter,[1] Kadi L. Saar,[1] Therese W. Herling,[1] Titus M. Franzmann,[3] Vasilis Kosmoliaptsis,[4,5,6] Simon Alberti,[3] F.-Ulrich Hartl,[7,8] Steven F. Lee,[2] Tuomas P.J. Knowles[1,9,*]

[1] Centre for Misfolding Diseases, Yusuf Hamied Department of Chemistry, Lensfield Road, Cambridge CB2 1EW, UK

[2] Yusuf Hamied Department of Chemistry, Lensfield Road, Cambridge CB2 1EW, UK

[3] Center for Molecular and Cellular Bioengineering, Biotechnology Center, Technische Universität Dresden, Tatzberg 47/49, 01307 Dresden, Germany

[4] Department of Surgery, University of Cambridge, Addenbrooke's Hospital, Hills Road, Cambridge CB2 0QQ, UK

[5] NIHR Blood and Transplant Research Unit in Organ Donation and Transplantation, University of Cambridge, Hills Road, Cambridge CB2 0QQ, UK

[6] NIHR Cambridge Biomedical Research Centre, Hills Road, Cambridge CB2 0QQ, UK

[7] Department of Cellular Biochemistry, Max-Planck Institute of Biochemistry, Am Klopferspitz 18, 82152, Martinsried, Germany

[8] Munich Cluster for Systems Neurology (SyNergy), Feodor-Lynen-Str. 17, 81377 Munich, Germany

[9] Cavendish Laboratory, Department of Physics, University of Cambridge, JJ Thomson Ave, Cambridge CB3 0HE, UK

[#] Contributed equally to this work

* Correspondence should be address to: tpjk2@cam.ac.uk

Abstract

Proteins constitute the molecular machinery of life and exert their biological function by interacting with other proteins, as well as by assembling into biomolecular complexes and higher order structures. Characterizing the sizes, interactions, and assembly states of proteins is thus key for understanding the normal functional behavior of proteins and for elucidating aberrant processes and interactions that can lead to dysfunction and disease. However, the physical characterization of proteins has remained a challenging problem due to the inherent compositional heterogeneity of protein mixtures as well as the polydisperse nature of protein complexes. Here, we address this challenge by demonstrating measurements of molecular diffusivity of single proteins and protein assemblies in microchannels using single-molecule fluorescence detection. The approach, termed single-molecule microfluidic diffusional sizing (smMDS), allows individual molecules to be counted directly, that is, in a digital manner, to enable single-molecule diffusional-sizing-based monitoring of protein hydrodynamic radii even within heterogenous multicomponent mixtures. Applying smMDS to a variety of protein systems, we show that the high sensitivity provided by smMDS enables ultrasensitive sizing of proteins down to the femtomolar concentration range. We further demonstrate the applicability of the approach towards affinity profiling of protein interactions at the single-molecule level and illustrate the potential of smMDS in resolving different assembly states of high- and low-molecular weight protein oligomers. Furthermore, we highlight the digital nature of the detection process by sizing multiple protein species within complex aggregation mixtures. Finally, we apply the approach to characterize nanoscale clusters of a phase separating protein system. Taken together, smMDS constitutes a versatile approach for digital, in-solution characterization of the sizes, interactions, and assembly states of proteins. We anticipate that smMDS will facilitate the discovery of new biomolecular mechanisms of proteins and will find broad applicability in the analysis of protein complexes in the biological, biophysical, and biomedical sciences, and beyond.

Introduction

Proteins form the molecular machinery of life and accomplish their biological function by interacting with other proteins as well as by assembling into biomolecular complexes and higher order structures.[1] The characterization of proteins is thus a key objective in many areas of biological and biomedical sciences, both for understanding the normal functional behavior of proteins and for elucidating aberrant processes and interactions that can lead to dysfunction and disease.[2–5] Proteins also serve as important therapeutic targets in drug discovery and clinical diagnostics, and are utilized as nanoscale building blocks in bionanotechnological applications.[6–8] A rigorous analysis of proteins and protein assemblies is therefore essential for advances in therapeutics development and the discovery of new biomaterials.[9–11] In particular features such as the molecular size of proteins and protein complexes, the strength of their interactions in terms of pairwise dissociation constants, and their assembly and oligomerization states are key parameters providing a heightened understanding of the biological function, malfunction, and design of proteins.[12–14] While such biophysical characterization of proteins has become routine, there still remain major hurdles that have not been addressed and which pose significant challenges for currently available biophysical approaches.

One particular challenge in the characterization of proteins lies in their inherent compositional complexity and heterogeneity. Protein systems often exist as mixtures of heterogeneous components, and exhibit polydispersity in terms of size and abundance. However, most biophysical approaches perform best when applied to pure homogeneous samples, and methods developed to enable the profiling of heterogeneous mixtures, such as gel filtration, electrophoresis, mass spectrometry, and surface plasmon resonance, pose significant problems.[15–18] These approaches are generally reliant on separation media, immobilization, or transferring the molecules from the solution phase into the gas phase, all of which may modify the distribution of sizes and thus make it challenging to relate the functional state of a protein to that probed under native solution conditions. While methods capable of studying molecules in bulk solution exist, such as analytical ultracentrifugation, isothermal titration calorimetry, and static and dynamic light scattering, they generally consume large amounts of protein and often require concentrations exceeding the biologically relevant range.[19–21] Many proteins are, however, present at low concentrations in biological media or samples are often available only in limited amounts. Therefore, ultrasensitive approaches that can size proteins and resolve heterogeneous mixtures of proteins at very low concentrations directly in solution are much sought after.

Micron-scale measurements of molecular diffusivity have proven to be a versatile and sensitive approach for probing the molecular sizes of proteins, their interactions, and assembly states in solution.[22–24] In particular, microfluidic diffusional sizing (MDS) has become an attractive, quantitative method for the characterization of proteins and protein complexes under native solution conditions.[25–35] MDS exploits the unique features of laminar flow in the microfluidic regime and measures the diffusive mass

transport of molecules across co-flowing sample and buffer streams within a microchannel.[24,25] By monitoring the diffusive spreading of analyte molecules at different downstream channel positions and analyzing the recorded diffusion profiles with advection–diffusion models, the sizes of analyte molecules can be quantified in terms of their hydrodynamic radii R_h. MDS can further probe the formation of biomolecular interactions and the assembly of analyte molecules into higher order structures by monitoring the increase in size associated with complex formation, and can retrieve binding affinities (i.e., dissociation constants K_Ds) through measurement of binding curves. However, despite the versatility of the approach, current implementations of MDS and other similar methods are limited in their ability to resolve compositional heterogeneities as their measured signals are based on ensemble readouts, which require deconvolution of an averaged signal in order to evaluate the complete size distribution of component species.[25,36,37] What exacerbates the situation is that diffusional analysis is often performed in a simplified manner, in which only relative amounts of analytes are measured. Such analysis only yields an average R_h for mixtures of differently sized species, thus limiting information content, in particular when studying heterogenous, multicomponent systems. Moreover, detection sensitivities for MDS and other sizing techniques are in the nanomolar to micromolar range, which hampers ultrasensitive protein detection at concentrations in the often desirable pico- to femtomolar range.[25,36,37]

Single-molecule detection methods, owing to their high sensitivity, bear great potential to overcome these challenges.[38,39] In particular, single-molecule fluorescence methods based on confocal detection are attractive readout modalities as they allow resolving distributions and heterogeneities underlying the ensemble without the necessity to deconvolve a signal, hence providing rich insights into subpopulations and compositional complexities.[40,41] Furthermore, they enable detection at ultra-low concentrations and provide direct digital readouts, that is, direct counting of individual proteins and protein complexes in solution.[42–46] However, direct readouts from single-molecule fluorescence cannot provide information on the sizes of molecules. Size information can only be retrieved indirectly, for example, through correlation analysis, as done in fluorescence correlation spectroscopy (FCS).[47,48] FCS and similar approaches, however, are reliant on calibration for absolute size determination and suffer from overparameterization, especially when heterogenous mixtures and polydisperse samples need to be resolved. Therefore, approaches that directly furnish single-molecule readouts with size information are much sought after. Accordingly, we reasoned that combining single-molecule fluorescence detection with diffusivity measurements of proteins in microchannels should lend itself a powerful platform to size proteins at the single-molecule level, enabling the characterization of heterogenous and multicomponent protein systems directly in solution.

Here, we demonstrate measurements of molecular diffusivity of single proteins and protein assemblies in microchannels using single-molecule fluorescence detection. By integrating a confocal fluorescence readout functionality into a microfluidic system that leverages the principles of MDS, we introduce an approach, termed single-molecule

MDS (smMDS), that enables diffusional-sizing-based monitoring of protein hydrodynamic radii at the single-molecule level. We first present the working principle and experimental implementation of the smMDS approach and show that, due to the digital nature of the detection process, it can be used to size single proteins and protein complexes with pico-to-femtomolar sensitivities. We further demonstrate the ability of smMDS to quantify the affinity of protein interactions at the single-molecule level and then show that smMDS enables resolving different assembly states of protein oligomers. We further use smMDS to characterize aggregate assemblies within multicomponent protein mixtures and apply the approach to the sizing of nanoscale clusters of a protein condensate system. Taken together, smMDS constitutes a versatile approach for digital, in-solution characterization of the sizes, interactions, and assembly states of proteins. We anticipate that smMDS will have widespread applicability in the biological and biomedical sciences for the discovery of novel biomolecular mechanisms underpinning protein function and malfunction. Furthermore, smMDS holds great potential for the analysis of protein interactions in drug discovery, clinical diagnostics and nanobiotechnological applications.

Results and Discussion

Working principle and experimental implementation of smMDS

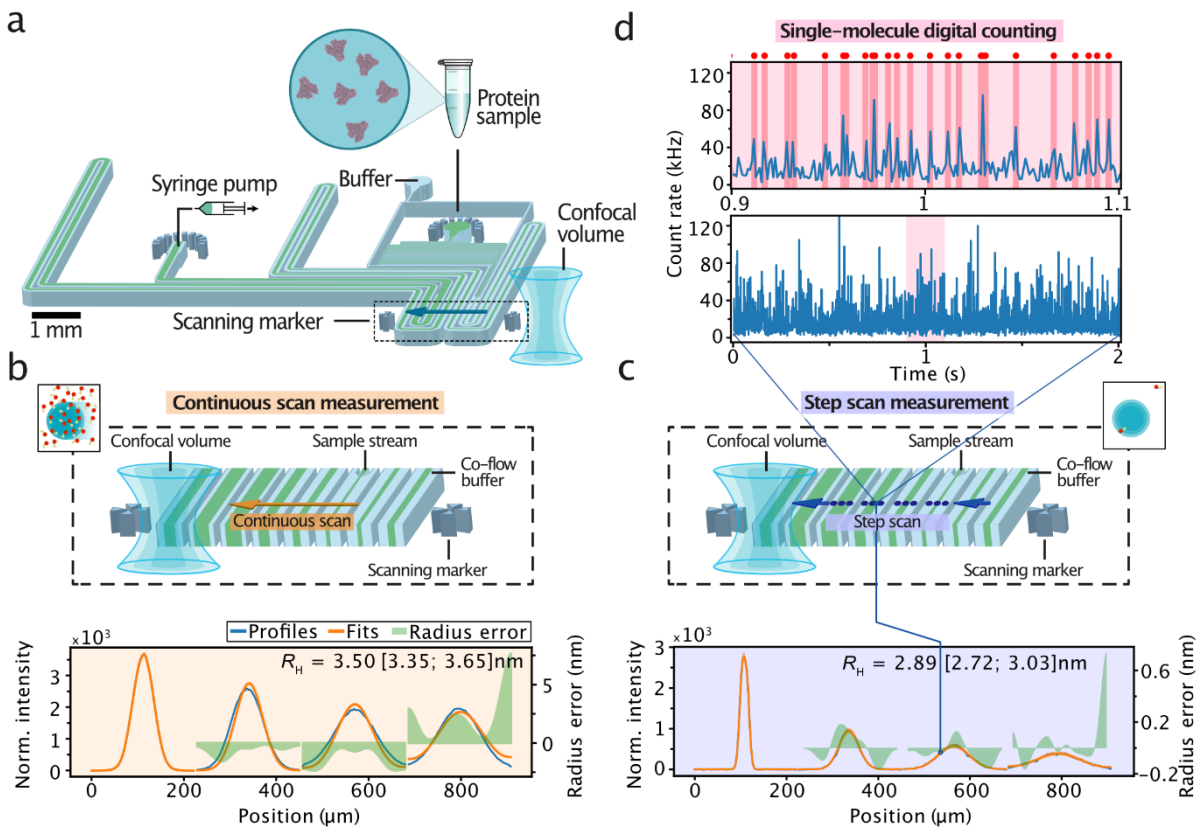

Figure 1. Working principle and experimental implementation of smMDS. (a) Schematic of the microfluidic chip design and integrated confocal scanning optics. The most relevant components are depicted. The dashed box highlights the scan region. The arrow indicates the scan trajectory across the four innermost channels. **(b)** Principle of continuous scan measurements. The confocal detection

volume is moved at a constant speed across the microfluidic device, enabling the recording of diffusion profiles from direct intensity readouts. This mode enables recording of diffusion profiles under ensemble conditions. An exemplary diffusion profile from a continuous scan measurement of human serum albumin (HSA) at 100 nM is shown. Diffusion profiles are shown as blue lines, experimental fits as orange lines, and errors as green bands. Extracted R_H [with errors] are given as insets. **(c)** Principle of step scan measurements. The confocal detection volume is moved in a stepwise manner across the device, collecting data at defined positions with each step for a certain period of time in the form of time traces (see panel d). This mode enables detection of individual molecules and the creation of diffusion profiles from single-molecule digital counting. An exemplary diffusion profile from a step scan measurement of α-synuclein at 10 pM is shown. **(d)** A single-molecule time trace (lower panel) as obtained from a step scan measurement is shown. The time trace in the upper panel is a zoom-in view of the red shaded area in the lower panel. Red dots and highlighting indicate bursts detected by the burst-search algorithm. The bin time is 1 ms in all traces.

Figure 1a illustrates the working principle and experimental implementation of smMDS. smMDS measures the molecular diffusivity of analyte molecules within a microfluidic chip. It operates based on the principles of MDS[25] and probes molecular diffusivity by flow-focusing an analyte stream between two auxiliary buffer streams within a microfluidic chip and then observing the diffusive spreading of analyte molecules to either side of the microfluidic channel as they travel downstream (see Methods). Because different positions along the channel correspond to different diffusion times, the tracking of the diffusive broadening of species at different channel positions allows quantification of the diffusion coefficient D and, thus, extraction of the size of analyte molecules in terms of R_h.[25] Experimentally, smMDS measurements are conducted by introducing sample containing fluorescent protein and buffer into the sizing chip and monitoring the micron-scale diffusive mass transport of molecules across the channel as they flow downstream the channel. Fluid flow in the channel is controlled by applying a negative pressure at the device outlet with a syringe pump (see Methods for details).

Detection in smMDS is done by a high sensitivity laser confocal fluorescence microscope functionality integrated into the microfluidic platform (see Methods). By scanning the confocal volume across the microfluidic chip at the mid-height of the channel perpendicular to the flow direction **(Figure 1a)**, fluorescence from passing analyte molecules is recorded. The scan trajectory is chosen such that various positions along the channel are probed, including positions that are close to the nozzle where the sample stream meets the co-flowing buffer medium, and others, further away, downstream of the channel. In our implementation, the four innermost channels of the device are scanned to obtain diffusion profiles, as these cover a wide range of distances and time points along the channel, such that biomolecular analytes with R_h-values on the 1–100 nm scale can be analyzed. The scan trajectory of the confocal volume in x,y,z-direction is set through two scan markers integrated within the microfluidic chip adjacent to the channels.

Scanning is conducted in two modes, by either continuously moving the confocal volume through the chip **(Figure 1b)**, or by moving the observation volume along the same trajectory in a stepwise manner, collecting data at defined positions with each

step **(Figure 1c)**. In continuous scan mode, diffusion profiles are rapidly acquired from direct intensity readouts. Hereby, the confocal volume is moved through the device at a constant scan speed (tens of µm/s) and the fluorescence intensity from analyte sample flowing through the confocal volume is recorded. This allows swift recordings of diffusion profiles under ensemble conditions, that is, at concentrations where many molecules are present in the confocal volume (i.e., typically at concentrations greater than tens of pM). An exemplary diffusion profile obtained from a continuous scan experiment is shown in **Figure 1b**. To extract R_H, the recorded diffusion profiles are analyzed using an advection-diffusion model (see Methods). Hereby, experimental profiles are fit to a set of simulated profiles obtained from numerical simulations. Using a least-squares error algorithm, the best matching simulated profile is identified via fitting to extract D and retrieve R_H using the Stokes-Einstein relationship **(Figure 1b)**.

In step scan mode, diffusion profiles are generated from time trace recordings along the scan trajectory. An exemplary diffusion profile obtained from a step scan experiment is shown in **Figure 1c**. Hereby, the confocal volume is parked at various positions along the trajectory and fluorescence signals from analyte flowing through the confocal volume are recorded for a certain period of time. Typically, between 200 to 400 scan steps across the chip are performed from start to end position and 2- to 4-second-long fluorescence traces are recorded at each position. Importantly, due to the high sensitivity of confocal detection, measurements in step scan mode enable the detection of individual molecules and, thus, the creation of diffusion profiles from single molecule counting **(Figure 1d)**. Hereby, bursts of fluorescence corresponding to the passage of single molecules through the confocal volume are recorded at each channel position. To estimate the number of molecules at each scanned position, a burst-analysis algorithm is employed (see Methods). This algorithm uses a combined maximum interphoton time (IPT_{max}) and minimum total number of photons (N_{min}) threshold criterium to extract single-molecule events from the recorded time trace at each position (see Methods). This approach has been shown to enable effective discrimination between photons that originate from single fluorescent molecules and those that correspond to background, thus allowing individual molecules to be counted directly, that is, in a digital manner.[46] From the detected number of molecules at each position, diffusion profiles are then created by plotting the number of counted molecules as a function of chip position. Extraction of R_H is done analogously as described above for continuous scan experiments by fitting the experimental diffusion profiles with our advection–diffusion analysis model. **Figure 1c** depicts fits and extracted R_H values for the example data set.

Sizing of proteins from bulk to single-molecule conditions by smMDS

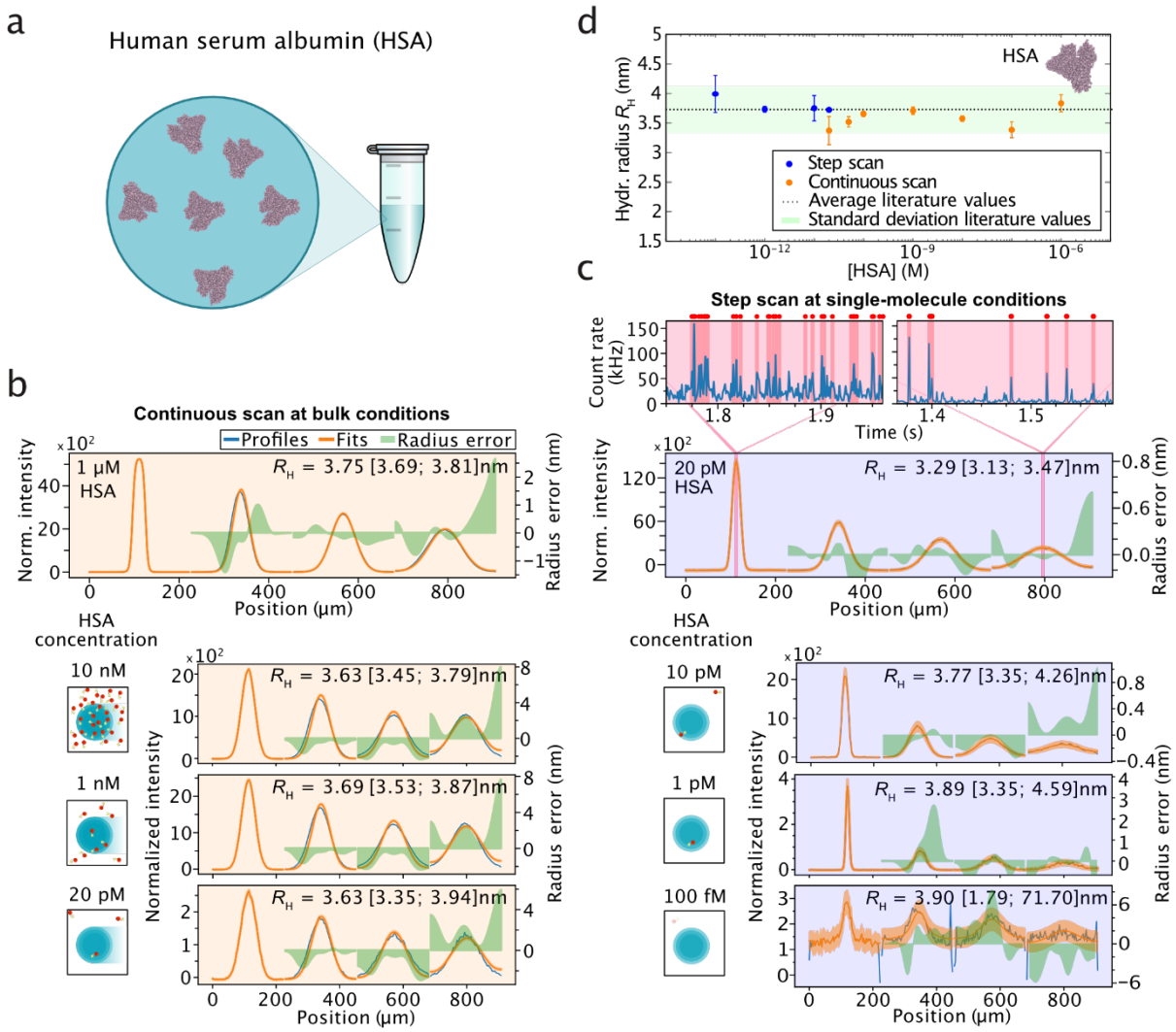

Figure 2. Sizing of proteins from bulk to single-molecule conditions by smMDS. (a) The sensitivity of smMDS and its capability to size proteins from bulk to single-molecule conditions was evaluated by measuring the size of HSA at varying protein concentrations. **(b)** Diffusion profiles for HSA as obtained from continuous scan measurements. From top to bottom: 1 µM, 10 nM, 1 nM, 20 pM HSA. Diffusion profiles are shown as blue lines, experimental fits as orange lines, and errors as green bands. Extracted hydrodynamic radii R_H [with errors] are given as insets. Schematics on the left depict the decrease in concentration. **(c)** Diffusion profiles for HSA obtained from step scan measurements under single-molecule conditions. From top to bottom: 20 pM, 10 pM, 1 pM, 100 fM HSA. The two highlighted plots on the top are exemplary single-molecule time trajectories recorded at two channel positions, as indicated. Red dots and highlighting indicate bursts detected by the burst-search algorithm. **(d)** R_H of HSA as obtained by continuous scan (orange points) and step scan (blue points) measurements. The dashed line indicates the average literature value for HSA (R_H = 3.73 ± 0.40 nm)[49–52] with the green band depicting the standard deviation of literature values.

In a first set of experiments, we sought to evaluate the sensitivity of smMDS and demonstrate its capability to determine protein size from bulk to single-molecule conditions. To this end, we labeled human serum albumin (HSA) with a fluorescent dye (Alexa 488) (see Methods) and performed concentration series measurements,

both in continuous and step scanning mode, by varying the HSA concentration. The recorded diffusion profiles of the series are shown in **Figure 2b,c**.

In the range from 1 µM down to tens of pM of HSA **(Figure 2b)**, sufficient molecular flux of HSA protein molecules allowed for the recording of diffusion profiles from continuous scan experiments. The obtained profiles show the characteristic broadening due to diffusion of molecules along the channels. Narrow peaks are observed at channel positions close to the nozzle where the sample meets the carrier medium, and peaks broadened as we probed further downstream channel positions. Modeling of the diffusion profiles using our advection–diffusion analysis approach yielded excellent fits. Extracted R_H values were amongst all concentrations, within error, in excellent agreement with previously reported values for HSA,[49–52] demonstrating the robustness and accuracy of the approach.

As we approached the single-molecule regime (i.e., at and below 20 pM) **(Figure 2c)**, we performed step scan measurements by moving the confocal in a stepwise manner across the channels. We observed bursts of fluorescence corresponding to the passage of single HSA molecules through the confocal volume **(Figure 2c, top panels)**. Using the burst-search algorithm (see Methods), we extracted single-molecule events from the recorded time traces at each position **(Figure 2c, top panels)** and created diffusion profiles by plotting the number of counted molecules as a function of chip position **(Figure 2c)**. In this way, we obtained diffusion profiles from digital counting for HSA from 20 pM down to 100 fM. Remarkably, also in this regime, by applying our advection–diffusion analysis, we obtained excellent fits of the experimental data and retrieved R_H values that were in agreement with previously reported values for HSA **(Figure 2d)**.[49–52] This shows that smMDS provides accurate size information over a broad range of concentrations and enables ultrasensitive sizing of proteins even in the picomolar to femtomolar concentration regime. We also evaluated the influence of parameter selection in the single-molecule regime on the extracted sizes **(Figure S1)** and found that a wide range of burst selection parameters, that is, varying thresholds for ITP_{max} and N_{min} (see Methods), yielded expected size information, supporting the robustness of the approach.

To compare the sensitivity of the smMDS technique to conventional MDS measurements, we conducted experiments utilizing fluorescence widefield imaging **(Figure S2)**. We performed concentration series measurements by varying the HSA concentration starting at 1 µM of labeled HSA and then gradually decreasing the protein concentration down to 100 and 50 nM. Image analysis yielded a clear profile only for the measurements at 1 µM and 100 nM protein concentration, and the expected size for HSA could only be recovered for the measurement at 1 µM protein. Notably, the measurement at 50 nM HSA yielded a featureless profile that could not be fitted and, hence, no R_H could be determined. This shows that conventional MDS experiments are limited to concentrations in the tens of nanomolar range. In comparison, the high sensitivity and digital detection capabilities afforded by smMDS allows measuring the size of proteins down to femtomolar concentrations, thereby

extending the sensitivity range of diffusional sizing experiments by more than five orders of magnitude.

Sizing of proteins and protein assemblies by smMDS

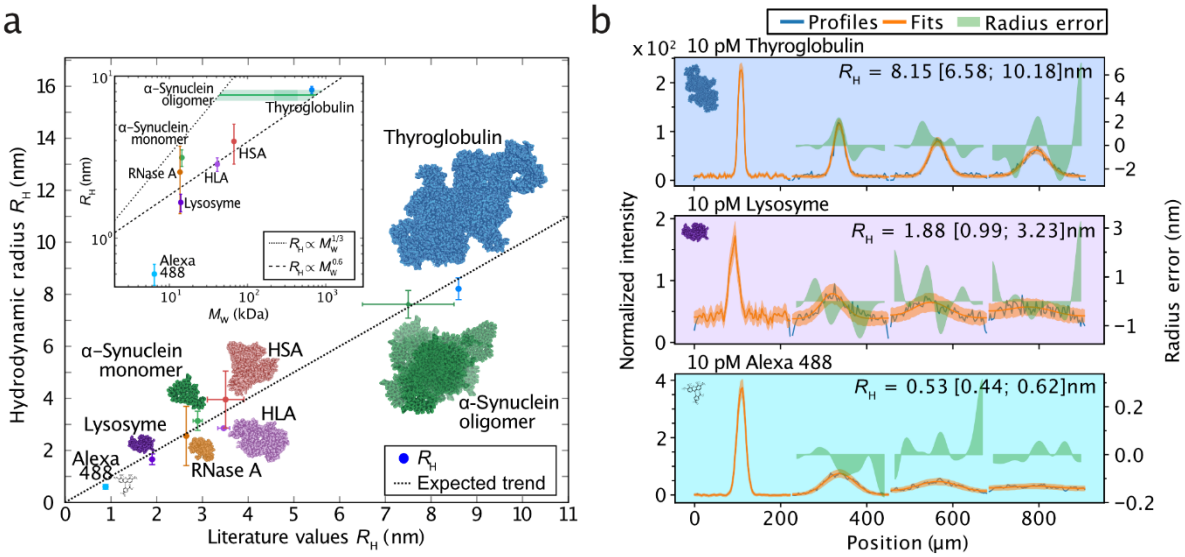

Figure 3. Sizing of proteins and protein assemblies by smMDS. (a) Experimentally determined R_H values as obtained from smMDS plotted against literature values. The dashed line depicts the expected trend. Data points were obtained from triplicate measurements at 10 pM sample concentration. Error bars denote standard deviations. Inset shows obtained R_H as a function of molecular weight M_W. The dashed and dotted lines denote scaling behavior of globular ($R_H \propto M_W^{1/3}$) and disordered ($R_H \propto M_W^{0.6}$) proteins, respectively. **(b)** Exemplary diffusion profiles for thyroglobulin (blue), HLA (violet), and Alexa 488 (cyan) at 10 pM sample concentration. Diffusion profiles are shown as blue lines, experimental fits as orange lines, and error as green bands. Extracted R_H [with errors] are given as insets.

Next, we sought to demonstrate the wide applicability of smMDS in determining the size of proteins and protein assemblies from single-molecule digital counting. We selected a set of analytes differing in size, including the proteins lysozyme, RNase A, α-synuclein, human leukocyte antigen (HLA), HSA, thyroglobulin, and oligomers formed by the protein α-synuclein. We also included the small organic fluorophore Alexa 488 in the series **(Figure 3a)**. The protein analytes were fluorescently labeled and purified before analysis (see Methods). We performed smMDS measurements at an analyte concentration of 10 pM and subjected the analytes to smMDS in step scan mode. We moved the confocal spot in a stepwise manner through the channels and extracted single-molecule events for each analyte at each channel position by digitally counting molecules to create diffusion profiles, which we fitted with our advection–diffusion model. Exemplary diffusion profiles for thyroglobulin, HLA, and Alexa 488 are shown in **Figure 3b**. The obtained R_H values from smMDS were then plotted against previously reported R_H values for Alexa 488,[53] lysozyme,[54] RNase A,[55] α-synuclein,[56] human leukocyte antigen (HLA),[57] HSA,[49–52] thyroglobulin,[58] and α-synuclein oligomers[59] **(Figure 3a)**. The values obtained by smMDS followed the expected trend within error. This demonstrates the excellent agreement between sizes obtained from smMDS and literature values, highlighting the reliability of the single-molecule

diffusivity measurements in size determination of protein analytes. In an additional analysis step, we plotted the experimentally obtained R_H values against the molecular weight M_W. Both, folded proteins (lysozyme, RNase A, HLA, HSA, thyroglobulin) and unfolded protein species (α-synuclein monomer and oligomers), followed the expected scaling behavior for globular ($R_H \propto M_W^{1/3}$) and disordered ($R_H \propto M_W^{0.6}$) proteins, respectively **(Figure 3a, inset)**.

Quantifying protein interactions by smMDS

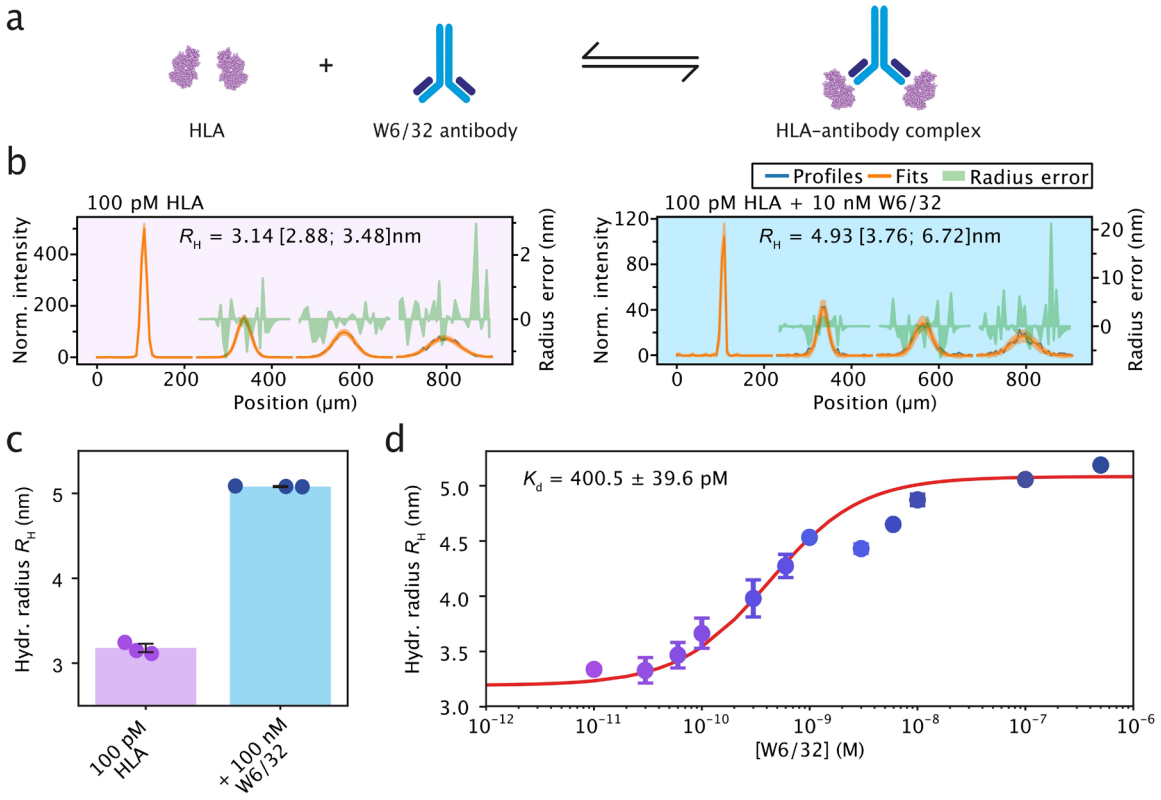

Figure 4. Quantifying protein interactions by smMDS. (a) Sizing and affinity measurement of an antibody–antigen complex by smMDS. Shown is a schematic for the binding interaction between the antigen HLA (A*03:01) and the HLA-antibody W6/32. **(b)** Exemplary diffusion profiles for the binding of the HLA-antibody W6/32 to HLA as obtained by step scan smMDS. The left panel shows a diffusion profile for 100 pM HLA. The right panel shows a diffusion profile for 100 pM HLA in the presence of 10 nM W6/32. Diffusion profiles are shown as blue lines, experimental fits as orange lines, and error as green bands. Extracted R_H [with errors] are given as insets. **(c)** Size increase upon complexation of HLA with W6/32 from R_H = 3.18 ± 0.04 nm for pure HLA to R_H = 5.08 ± 0.01 nm for the complex in the presence of 100 nM W6/32. **(d)** Binding isotherm obtained from a titration of 100 pM HLA with increasing concentrations of W6/32. For analysis, the binding isotherm was fitted with a binding model assuming two antigen molecules binding one antibody.[57] The dissociation constant was found to be K_d = 400.5 ± 39.6 pM. Error bars are standard deviations from triplicate measurements.

Next, we set out to demonstrate the capability of smMDS in determining the affinity of biomolecular interactions at the single-molecule level. Interactions of proteins with secondary biomolecules, in particular with other proteins, are of great importance across the biosciences, and quantitative measurements of affinity constants in the form of K_Ds have become vital in biomedical research and clinical diagnostics, for

example, for affinity profiling.[29,57,60,61] Diffusional sizing allows for the detection of biomolecular interactions by monitoring the increase in size associated with binding and complex formation.[25,31] By acquiring binding isotherms, affinity constants of the interaction can be determined in solution, without the need for purification or for immobilization on a surface. So far, diffusional sizing has been limited to the sizing and quantification of protein interactions at bulk nanomolar concentration levels—with smMDS, this barrier can be overcome.

To demonstrate the detection of biomolecular interactions and quantification of binding affinities by smMDS in a digital manner, we probed binding of a clinically relevant antibody–antigen interaction. Specifically, we investigated the binding interaction between HLA A*03:01, an isoform of the major histocompatibility complex type I (MHC) and a key factor in the human immune system,[62] and the antibody W6/32, an antibody that binds to all class I HLA molecules **(Figure 4a)**.[63] We performed a series of step scan smMDS experiments by titrating HLA antigen (labeled with Alexa 488), at a constant concentration of 100 pM, with increasing amounts of the unlabeled W6/32 antibody. Exemplary diffusion profiles for pure HLA at 100 pM, and 100 pM HLA titrated with 10 nM of W6/32 are shown in **Figure 4b**. smMDS diffusion profiles, from three repeats, were acquired and fitted to obtain effective R_H across the concentration series. We observed an increase in average hydrodynamic radius from R_H = 3.18 ± 0.04 nm for pure HLA, corresponding to a molecular weight of 50 kDa, as expected for HLA, to R_H = 5.08 ± 0.01 nm for the saturated complex, corresponding to a molecular weight of 215 kDa, consistent with the binding of a 150 kDa antibody to HLA **(Figure 4c)**. By fitting the binding isotherm **(Figure 4d)**, we determined the dissociation constant to be K_d = 400.5 ± 39.6 pM, consistent with previous results.[57] Importantly, HLA is an extensively used clinical biomarker to assess, for example, the risk of allograft rejection, and typically present at low concentrations. These results here therefore outline a path towards detection and affinity profiling of antibody responses when only low concentrations of samples are available. More generally, these results highlight the possibility to quantify biomolecular interactions from single-molecule digital measurements at ultralow concentrations.

Resolving protein oligomeric states by smMDS

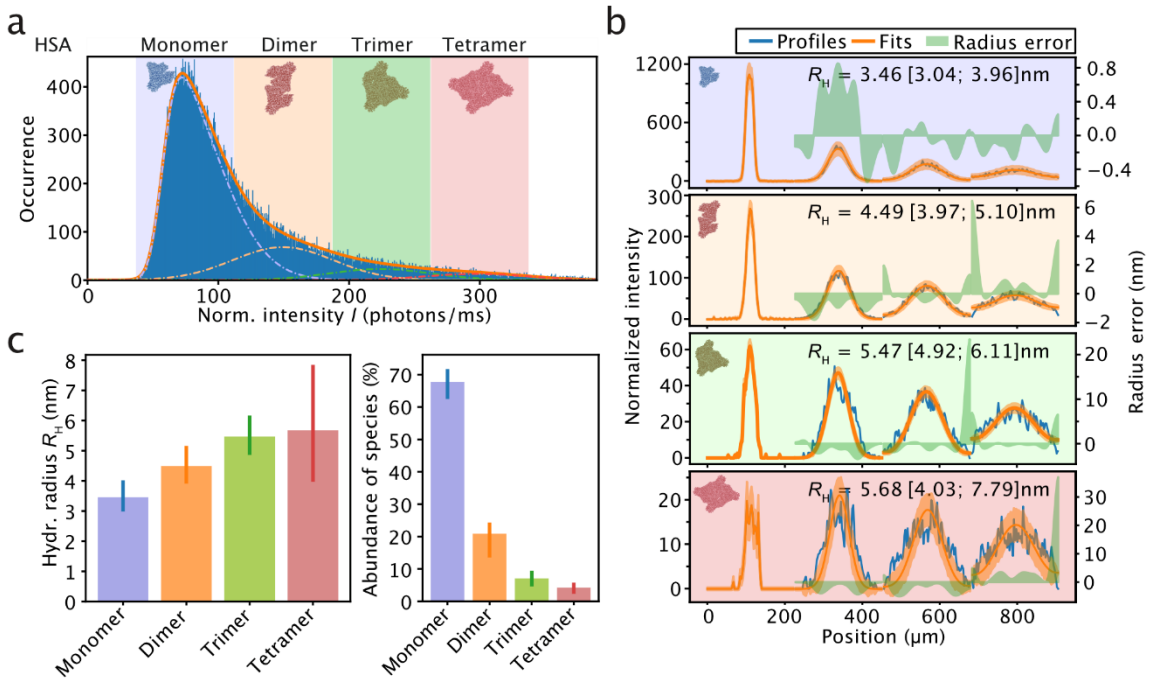

Figure 5. Resolving protein oligomeric states by smMDS. (a) Sizing of low-molecular weight oligomers formed by the protein HSA. Shown is a burst intensity histogram, which displays all bursts extracted from a smMDS step scan measurement of HSA (10 pM). Intensities are normalized intensities with respect to burst duration. Four regions (blue, orange, green, and red), which correspond to HSA monomer and HSA dimer, trimer and tetramer, are defined in the burst intensity histogram. These were obtained by fitting the distribution with an asymmetric Gaussian for monomeric HSA and three Gaussian functions for dimeric, trimeric, and tetrameric HSA (see main text for details). The center positions for the oligomers (dimer: 150.66 photons/ms, trimer: 225.99 photons/ms, and tetramer: 301.32 photons/ms) are multiples of the normalized intensity of the monomer ($I_{monomer}$ = 75.33 photons/ms). The widths of the regions reflect one standard deviation of the distribution of monomeric HSA ($\sigma_{monomer}$ = 37.44 photons/ms). **(b)** Diffusion profiles for HSA monomer, dimer, trimer and tetramer generated from bursts within each of the four regions in the burst intensity histogram (panels are color coded according to the colors used in panel a). Each profile was fitted to extract size information. Diffusion profiles are shown as blue lines, experimental fits as orange lines, and error as green bands. Extracted R_H [with errors] are given as insets. **(c)** Species-resolved R_H (left panel) and abundance of HSA oligomers (right panel). R_H were obtained from diffusion profile fits shown in panel b. Abundance was obtained from Gaussian fitting in panel a. The error bars correspond to the 99% percentile confidence intervals obtained from bootstrapping.

Many proteins fulfill their biological roles not as monomeric species but as oligomeric assemblies, which often exhibit significant heterogeneity in terms of their degree of oligomerization and relative abundance.[64–66] Oligomeric forms of proteins play important functional roles in cellular physiology but are also implicated in diseases such as neurodegeneration.[67–69] Resolving the degree of oligomerization and thus the size of such heterogenous protein populations is, however, extremely challenging with currently available biophysical techniques. A key feature of smMDS is that it has the capability to directly distinguish between various assembly states of a protein based on a difference in their emitted fluorescence signals. This feature is afforded by the single-molecule sensitivity of smMDS and enables the creation of diffusion profiles

from subspecies that make up the heterogeneous population. To demonstrate this capability, we set out experiments with two protein oligomer systems that are inherently heterogeneous and have distinct functions in biology and disease.

In a first set of experiments, we determined the sizes of low-molecular weight oligomers formed by the protein HSA **(Figure 5)**. Serum albumins are known to exist in an equilibrium of monomers, dimers, trimers, and tetramers, which are populated with decreasing relative abundance.[70] At the single-molecule level, such different oligomeric states can be discriminated through brightness analysis of fluorescence bursts.[71,72] In this analysis, different species can be distinguished based on their emitted fluorescence intensity because the magnitude of the observed intensity scales directly with the number of individually dye-labeled monomer units present in an oligomeric assembly. By applying differential thresholding, oligomeric states can then be discriminated, which provides an opportunity for smMDS to create distinct diffusion profiles for each oligomeric subspecies, enabling their independent species-specific sizing. To demonstrate this, we set out smMDS measurements to resolve the sizes of HSA monomers, dimers, trimers, and tetramers. We subjected labeled HSA to smMDS at 10 pM protein concentration and performed step scan measurements to extract single-molecule events from single-molecule time trace recordings. We then displayed the extracted normalized burst intensities from all recorded burst events of the measurement in a burst intensity histogram **(Figure 5a)**. This allowed us to display single-molecule burst events according to their brightness and assign regions of intensity for the monomeric and the different oligomeric HSA species. Accordingly, the main peak in the histogram reflects the average intensity of monomeric HSA. We extracted this intensity by fitting the distribution with an asymmetric Gaussian function, which reflects the skewedness of the burst intensity distribution due to the undersampling effects at short burst times, and retrieved a mean intensity for the monomer of $I_{monomer}$ = 75.33 photons/ms with a standard deviation $\sigma_{monomer}$ of 37.44 photons/ms. Since oligomers contain as many fluorophores as monomer units, dimeric, trimeric, and tetrameric forms of HSA emit at multiples of the normalized intensity of monomeric HSA, due to the increasing number of fluorophore present within the assembled states. We therefore defined regions at two-, three-, and four-fold of the normalized intensity of the monomer, corresponding to HSA dimer, trimer, and tetramer, respectively by fitting the burst intensity distribution with three Gaussian functions. The widths of these oligomer regions were assumed to have the same standard deviation as the monomeric protein. The resulting fit for all Gaussians, including the asymmetric Gaussian for the monomer, described well the experimental burst intensity distribution. We then generated diffusion profiles from the bursts within each of the four regions and fitted the profiles to extract size information from the respective monomer/oligomer range **(Figure 5b)**. The extracted sizes of the four different regions correspond to molecular weights of proteins of 65 (44–99) kDa, 145 (99–214) kDa, 266 (192–370) kDa, and 298 (192–790) kDa, respectively, and thus are, within error, in very good agreement with the sizes of monomeric, dimeric, trimeric, and tetrameric HSA (66 kDa, 144 kDa, 199 kDa, 266 kDa, respectively). In

addition to subspecies-resolved sizing of oligomers, our analysis can also provide information on the relative abundance of oligomeric species. The areas under the curves, as obtained from Gaussian fitting of the burst intensity histogram, reflect the abundance of oligomeric species. We obtained relative abundances of 67.9% for the monomer, 20.7% for the dimer, 7.2% for the trimer, 4.2% for the tetramer **(Figure 5c).** These values are in very good agreement with the expected equilibrium distribution of serum albumins.[70] Overall, our analysis here shows that smMDS can afford species-resolved insights into oligomer size and abundance.

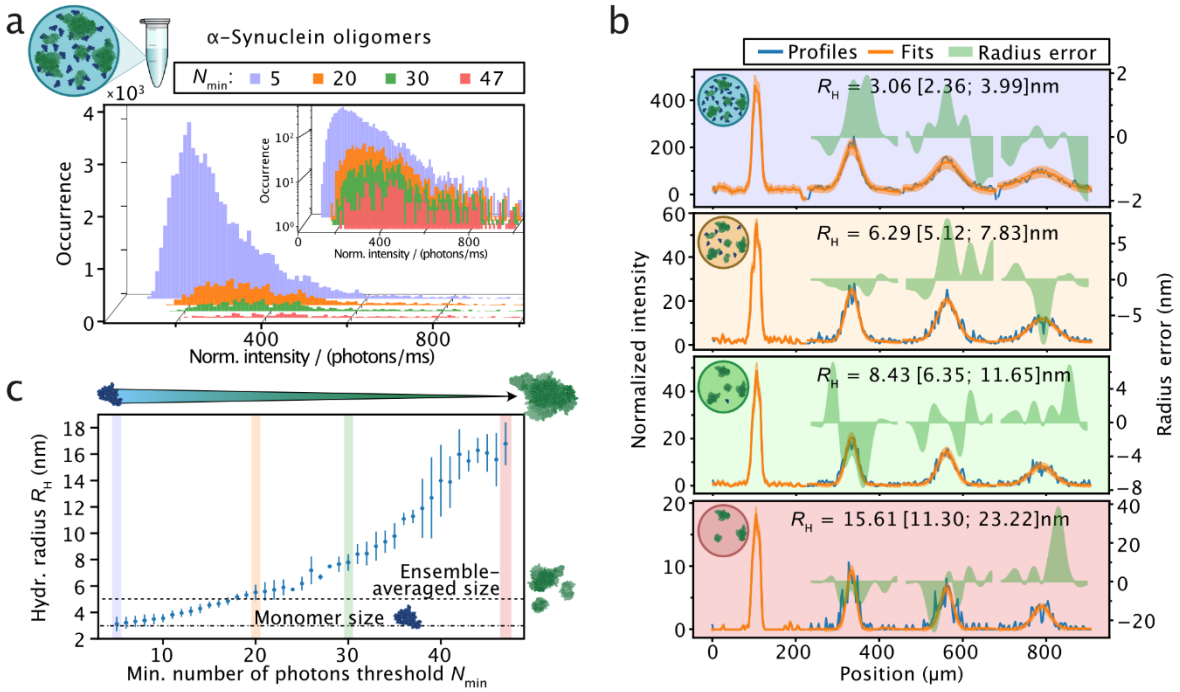

Figure 6. Sizing of heterogenous oligomer populations by smMDS. (a) A heterogenous mixture of α-synuclein oligomers (10 pM) was analyzed by step scan smMDS and single-molecule burst events were extracted using the burst search algorithm. To differentiate between differently sized assembly states of α-synuclein oligomers, the value for the minimum number of fluorescence photons threshold was varied, while keeping the inter-photon time threshold constant. This allowed for the creation of burst intensity distributions, which differ in molecular brightness. Exemplary burst intensity distributions for four different threshold values are shown (light blue: 5 photons, orange: 20 photons, green: 30 photons, red: 47 photons). The inset displays burst intensity histograms in semi-log scale. Intensities are normalized intensities with respect to burst duration. **(b)** Exemplary diffusion profiles generated from burst intensity distributions with four different minimum number of fluorescence photons threshold values (panels are color coded according to the specific threshold values used in panel a). Diffusion profiles are shown as blue lines, experimental fits as orange lines, and error as green bands. Extracted R_H [with errors] are given as insets. **(c)** Extracted R_H of oligomer subspecies displayed as a function of the different minimum number of photon threshold values used in the burst search algorithm. The colored vertical bars indicate threshold values used in panel a and b. The horizonal dashed line indicates the ensemble-averaged size of the entire oligomer population, generated from a diffusion profile from all bursts obtained from the measurement. R_H of monomeric α-synuclein, as measured by smMDS, is indicated as a dotted horizontal line.

In a further set of experiments, we analyzed a heterogenous mixture of α-synuclein oligomers **(Figure 6)**. Oligomeric forms of the protein α-synuclein are considered to

be central to the pathology of Parkinson's disease and hallmarked by a high degree of heterogeneity in terms of size and structure.[73–75] Their characterization is an area of intense interest, not least because such information is useful in drug development activities, however, tools that can directly resolve the heterogeneity of these nanoscale assemblies in solution are scarce.[59,76–80] To address this challenge and characterize structural heterogeneity of α-synuclein oligomers, we analyzed a heterogenous mixture of α-synuclein oligomers by smMDS. We injected oligomers produced by lyophilization of Alexa 488-labeled α-synuclein into the microfluidic sizing chip at a concentration of 10 pM and performed step scan measurements to digitally extract single-molecule events of passing α-synuclein oligomer molecules at each channel position. To create diffusion profiles from subspecies, we selected bursts with different fluorescence intensities to resolve differently sized assembly states of oligomers within the mixture. Here we took an alternative approach as compared to the analysis of HSA oligomers presented above and extracted bursts by varying the minimum number of fluorescence photons in the burst search algorithm, while keeping the inter-photon time threshold constant. This allowed us to effectively differentiate between single-molecule burst events that differ in their molecular brightness (see burst intensity histograms shown in **Figure 6a**). In this way, diffusion profiles from assemblies that differ in their fluorescence intensity and, hence, size in terms of R_H were generated. Exemplary diffusion profiles for four different thresholds are shown in **Figure 6b.** These profiles were then fitted with our advection–diffusion model to extract size information. We applied this analysis to a range of photon thresholds (i.e., 5–50 total number of photons) and generated a plot of the extracted sizes versus photon thresholds from the diffusion profile series **(Figure 6c)**. The measured sizes span from R_H = 3.6 nm for the smallest photon threshold value to R_H = 16.5 nm for the largest threshold value. The value at the smallest threshold reflects the size of α-synuclein monomers (R_H = 3.1 ± 0.4 nm) **(c.f. Figure 3)**, while higher values reflect α-synuclein oligomer subpopulations, spanning a range of R_H values greater than 10 nm. Notably, the R_H-value distribution obtained by smMDS is in excellent agreement with a previous study,[76] which described the variability in size of monomer/oligomers to be in the range of 3–16 nm. In comparison, we also generated a diffusion profile from the entire set of bursts detected at each position, which through fitting yielded an ensemble-averaged value of the size of the oligomer population (R_H = 5.2 ± 0.1 nm) **(Figure 6c)**. To complement the analysis demonstrated here based on varying the minimum number of fluorescence photon thresholds in the burst search algorithm, we also performed size analysis of α-synuclein oligomers by selecting defined regions in the burst intensity histogram as done for HSA. The results are shown in **Figure S3** and are in very good agreement with the results obtained by varying the fluorescence photon thresholds value **(c.f. Figure 6c)**. Taken together, our analyses here demonstrate the versatility of smMDS in resolving the size distributions of a heterogenous oligomeric protein samples.

Sizing of multiple species within a heterogenous aggregation mixture by smMDS

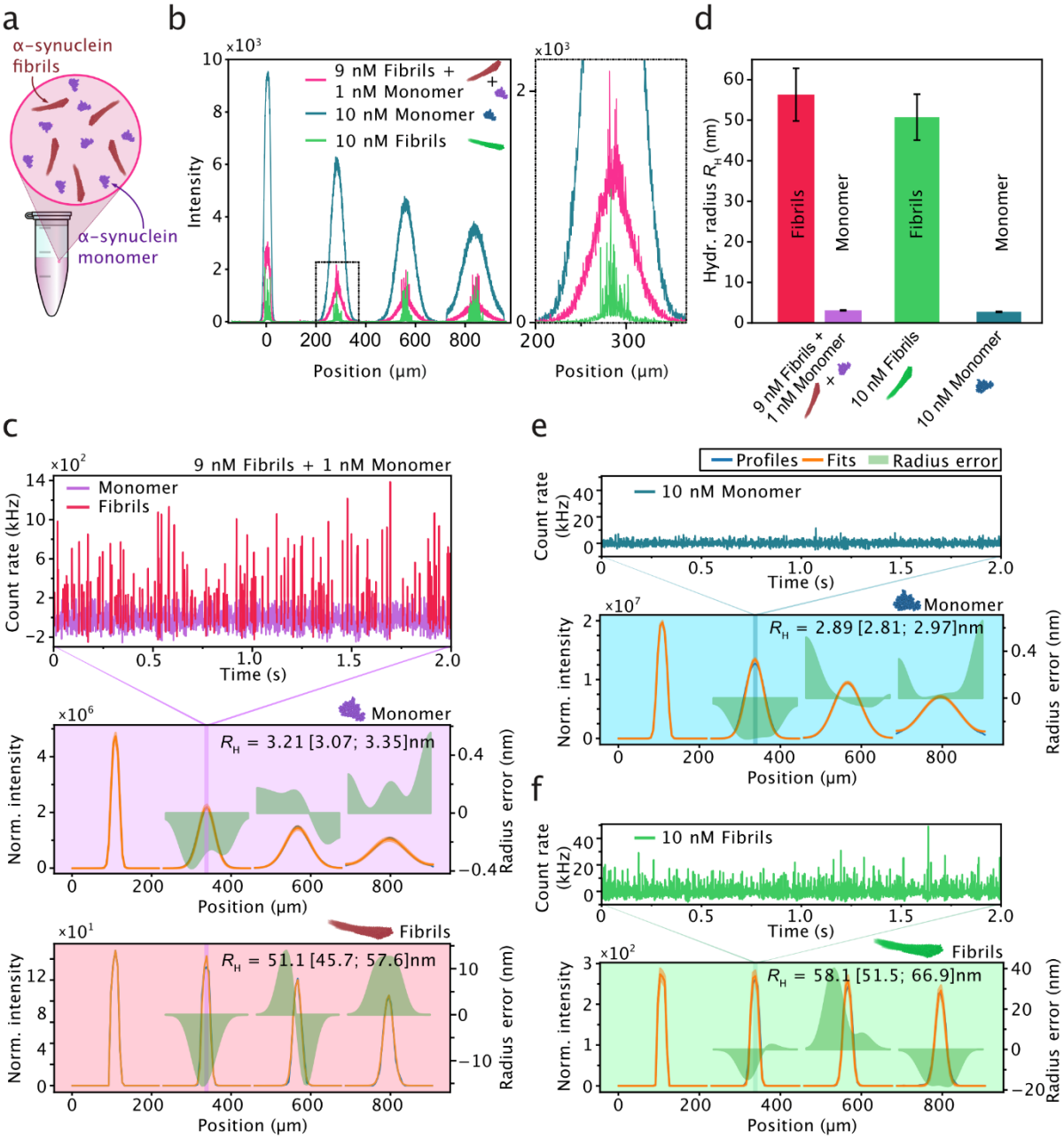

Figure 7. Sizing of multiple species within a heterogenous aggregation mixture by smMDS. (a) Schematic of an aggregation reaction composed of monomeric α-synuclein protein and large fibrillar species. **(b)** Sizing of α-synuclein fibrils in the presence of an excess of monomeric α-synuclein. Continuous scan diffusion profiles (left panel) for pure monomeric α-synuclein (10 nM) (blue), α-synuclein fibrils (10 nM monomer equivalent) (green), and a mixture of α-synuclein fibrils (9 nM monomer equivalent) and monomeric α-synuclein (1 nM) (pink). The right panel is a zoom-in as indicated by a dashed box in the left panel. Bursts correspond to the passing of single α-synuclein fibrils through the confocal detection volume. **(c)** Step scan measurement of a mixture of α-synuclein fibrils (9 nM monomer equivalent) and monomeric α-synuclein (1 nM). The top panel shows an exemplary fluorescence time trace (1-ms binning) at diffusion profile position 340 μm, as indicated. An intensity threshold was applied to separate signal stemming from fibrils (red) and monomer (purple). The bottom panels show diffusion profiles created from the fibril and monomer signals, respectively. The fibril signal

was obtained by intensity thresholding at 250 kHz after applying a Savitzky-Golay smoothing filter. The remaining signal formed the monomer signal. Diffusion profiles are shown as blue lines, experimental fits as orange lines, and errors as green bands. Extracted hydrodynamic radii R_H [with errors] are given as insets. **(d)** Comparison of extracted sizes from triplicate step scan measurements. Shown are R_H of species extracted from a mixture of α-synuclein fibrils (9 nM monomer equivalent) and 1 nM monomeric α-synuclein (red and purple, respectively), pure monomeric α-synuclein (blue), 10 nM monomer equivalent of α-synuclein fibrils (green). The error bars correspond to the standard deviation of triplicate measurements. **(e)** Step scan measurement of pure α-synuclein (10 nM). The top panel shows an exemplary fluorescence time trace (1-ms binning) at diffusion profile position 338 µm, as indicated. **(f)** Step scan measurement of pure fibrils (10 nM monomer equivalent). The top panel shows an exemplary fluorescence time trace (1-ms binning) at diffusion profile position 340 µm, as indicated.

Many protein systems are heterogeneous, multicomponent mixtures consisting of proteins and protein assemblies that differ in size by several orders of magnitude. For example, aggregation mixtures are made up of monomeric protein and large fibrillar species.[3,30] Often, one of the components (e.g., the monomeric protein) is present in large excess, while the other one (e.g., fibrillar species) is only present in small amounts. Approaches that can quantify the sizes of such differently populated species are much sought after, yet lacking. smMDS can fill this gap as it has the capability to size molecules and assemblies in heterogeneous mixtures even when an excess of one of the molecular species is present at bulk levels.

To illustrate the potential of sizing mixtures by smMDS that are compositionally heterogenous, we set out experiments with a sample system composed of fibrils formed by the protein α-synuclein, a key component in the pathology of Parkinson's disease,[81,82] and an excess of monomeric α-synuclein at nanomolar concentrations **(Figure 7a)**. Such a mixture is often encountered in assays that probe the mechanisms underlying protein aggregation and amyloid formation.[30,83] We first performed continuous scanning experiments on pure α-synuclein fibrils (at 10 nM monomer equivalent concentration) and pure α-synuclein monomer (at 10 nM concentration) to establish the signature of the two species **(Figure 7b)**. For the fibril-only sample **(Figure 7b, green profile)**, we observed burst events of high fluorescence intensity that were narrowly distributed around the center of the channels. The high burst intensity stems from the large number of fluorophores that are contained in a single fibril (>10% of the monomers are fluorescently labelled within fibrils, see Methods). The narrow distribution of bursts located at the center of the channel indicates a low diffusion coefficient and correspondingly a large size, as expected for fibrillar aggregates. For the monomer sample **(Figure 7b, blue profile)**, the sizing profiles exhibited a broader spread, due to the larger diffusivity of the monomeric units as compared to fibrils. The monomer signal is continuous because nanomolar concentrations are used, and therefore multiple monomeric units traverse the confocal detection volume at the same time, resulting in a bulk fluorescence signal rather than individual single-molecule events, as observed for the fibril sample. In addition to establishing the signatures of fibrillar and monomeric samples, we probed a sample mix containing α-synuclein fibrils and an excess of the monomeric protein **(Figure 7b, pink profile)**. The diffusion profile of the mixture now exhibited characteristic

signatures for both fibrils and monomeric protein, with broadened fluorescence at the profile base, reflecting monomeric protein, in addition to bright bursts on top of the monomeric signal that were narrowly distributed at the center of the channel, reflecting signals from fibrils.

To demonstrate that smMDS is able to size both the monomeric subpopulation present at bulk levels and the fibrils present at single-particle concentrations, we performed step scan measurements with a mix containing α-synuclein fibrils and an excess of the monomeric protein **(Figure 7c)**. An example fluorescence time trace is shown in **Figure 7c (top panel)**. Fibrils are clearly detectable as bursts above the mean signal, which corresponds to the bulk monomer signal. From these traces, we separated the bulk monomer signal from the fibril burst signals by intensity thresholding. Specifically, fibrils were detected as bursts that exhibit a fluorescence count rate of >250 kHz, after applying a Savitzky-Golay smoothing filter. The remaining signal (i.e., the mean bulk signal in the fluorescence time traces in the absence of fibril signal) formed the signal for the monomer. From the extracted fibril and monomer signal, we created diffusion profiles for the two species **(Figure 7c, bottom panels)** and subjected these profiles to fitting using our advection–diffusion model to extract size information of the two species. The sizes of monomer and fibrils species, from triplicate measurements **(Figure 7d)**, were estimated to be $R_{H,monomer}$= 3.23 ± 0.04 nm and $R_{H,fibrils}$ = 56.43 ± 6.69 nm. As a control, we also performed step scan measurements for the fibril-only and monomer-only sample for comparison **(Figure 7e and f, respectively)** and obtained sizes which were, within error, in excellent agreement with the ones obtained from the sample mix, thereby validating our approach **(Figure 7d)**.

Together, these experiments here show that smMDS has the capability to quantify differently sized molecules or assembly states of a protein within a heterogeneous mixture, even when an excess of one of the molecular species is present at bulk levels. These findings are significant as such an approach allows for the simultaneous probing of differently populated species, for example, in kinetic protein misfolding and aggregation studies.

Sizing of nanoscale clusters by smMDS

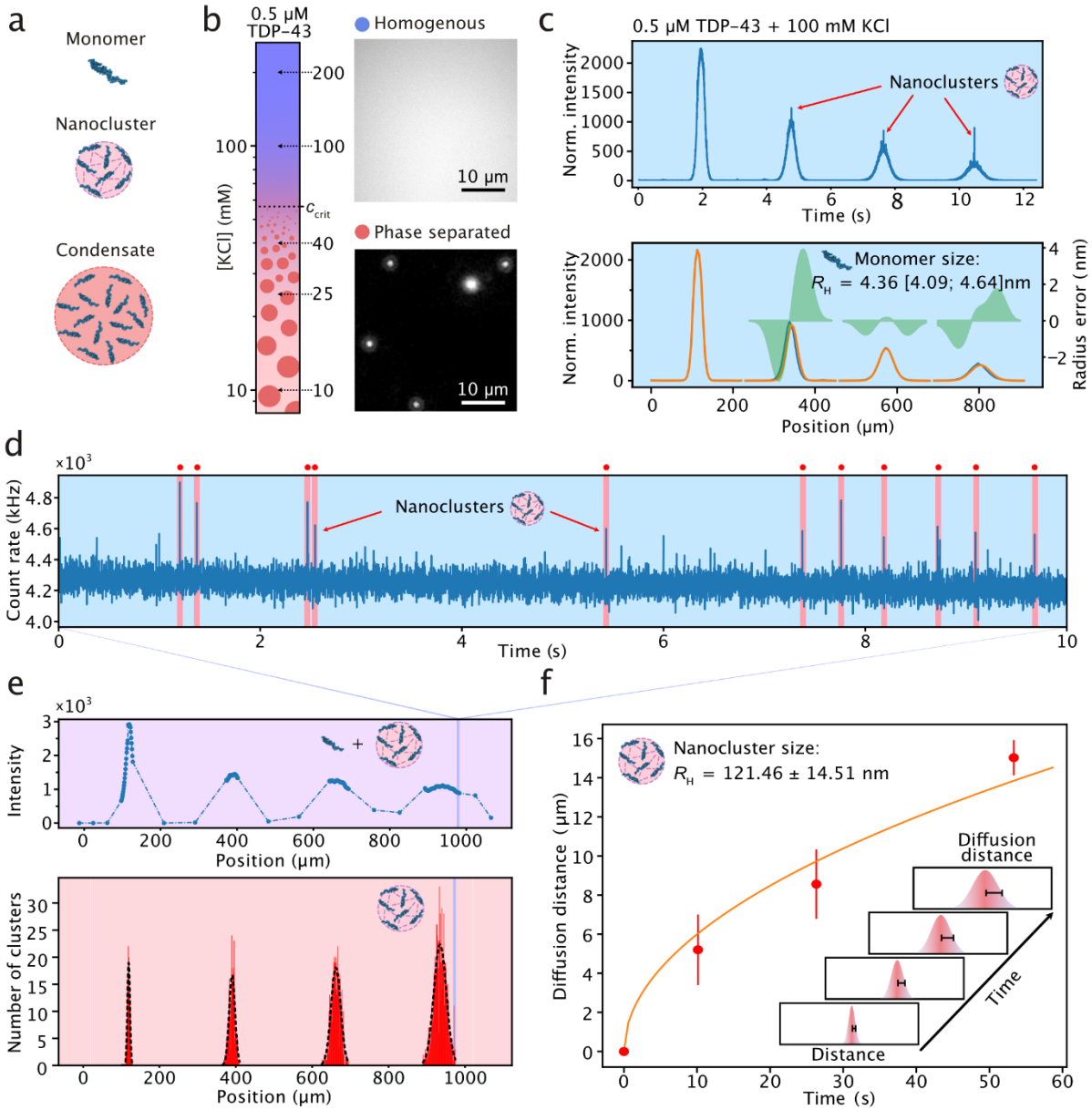

Figure 8. Sizing of nanoscale clusters by smMDS. (a) Schematic of TDP-43 phase separation and the formation of nanoscale clusters in a pre-phase separating regime. **(b)** Phase separation behavior of GFP-tagged TDP-43 (0.5 μM) as a function of KCl concentration as observed by conventional microscopy imaging. Microscopically visible condensates (1–2 μm in diameter) appear below a critical KCl concentration (c_{crit}) of ~50 mM, and condensates decrease in size as the critical concentration is approached. No microscopically visible condensates are detectable above c_{crit} and the solution appears clear and well-mixed. The phase diagram on the left was generated from measurements at five KCl concentrations and a protein concentration of 0.5 μM. Representative widefield fluorescence microscopy images at 100 mM and 25 mM KCl are shown on the right. **(c)** Continuous scan diffusion profiles for 0.5 μM GFP-tagged TDP-43 at 100 mM KCl. The upper panel shows the diffusion profile as obtained from the continuous scan measurement directly. Diffusive broadening along the channel characteristic of monomeric protein is observed. In addition, bright bursts that are narrowly distributed at the center of the channels are visible. These bursts correspond to the passing of nanoclusters through the confocal detection volume. The bottom panel is a re-binned diffusion profile from the upper panel to extract the size of monomeric GFP-tagged TDP-43. Diffusion profiles are shown as blue lines,

experimental fits as orange lines, and error as green bands. Extracted R_H [with errors] are given as insets. **(d)** Exemplary fluorescence time trace (1-ms binning) from a step scan measurement at diffusion profile position 960 µm, as indicated in panel e. Nanoclusters were detected as bursts that exhibit a fluorescence signal >5 standard deviations above the mean. Detection events are indicated with red highlighting and red dots. **(e)** Total intensity of a segmented step scan across the chip (top panel) and histogram of detected nanocluster events as a function of chip position (bottom panel). 1-µm step scans were performed in the center regions of the channels, whereas 50-µm step scans were performed outside these regions. Gaussian distributions were fit to each peak in the histogram to extract a mean diffusion distance at each channel position. The four channel positions correspond to 0, 10, 26, and 55 seconds of travel within the channel. **(f)** Plot of mean diffusion distance versus time of travel within the channel. The inset describes graphically how the diffusion distances were determined. The diffusion distance corresponds to the half of the full-width half maximum (FWHM) of the Gaussian distributions at each measurement point. The half of the FWHM of the Gaussian function from time point zero was used for normalization and subtraction from all other time points. Data points are from three repeats and the error bars correspond to the standard deviation of the repeats. The orange line shows the fit according to Eq. 1. The extracted R_H of TDP-43 nanoclusters is given as an inset.

In a final set of experiments, we applied the smMDS approach to the characterization of nanometer-sized clusters of a phase separating protein system. Biomolecular condensates **(Figure 8a)** formed through phase separation are important players in cellular physiology and disease[84,85], and emerge from the demixing of a solution into a condensed phase and a well-mixed phase.[86,87] Condensates typically have sizes in the micrometer range and, as such, are typically observable by conventional microscopy imaging.[88–90] Recent evidence, however, suggests that phase separation-prone proteins, such as the DNA/RNA binding protein fused in sarcoma (FUS), can also form nanoscale assemblies **(Figure 8a)**, well-below the critical concentration at which phase separation occurs (i.e., pre-phase separating regime).[91–94] These so-called nanoscale clusters have sizes in the tens to hundreds nanometer regime, and thus are beyond the resolution of conventional optical imaging systems. Moreover, as these species are low in abundance and present in a high background of dilute phase protein concentration, they are typically hard to detect by classical approaches. Here, we demonstrate that smMDS can determine the size of nanoscale assemblies formed by the protein TAR DNA binding protein 43 (TDP-43) at sub-saturating concentrations. We further show that smMDS allows for the determination of nanocluster abundance and composition.

First, we mapped out a one-dimensional phase diagram of the protein TDP-43 **(Figure 8b)** to assess the phase separation behavior of the protein with respect to changes in salt concentration. GFP-tagged TDP-43 at 0.5 µM protein concentration formed microscopically visible condensates below a critical salt concentration c_{crit} of 50 mM KCl. No condensates were visible by conventional fluorescence microscopy above that salt concentration and the solution appeared clear and well-mixed. Next, in order to assess whether TDP-43 forms nanoscale assemblies, we performed smMDS measurements at conditions where no microscopically visible condensates could be detected (i.e., well above c_{crit}). To this end, we first performed a continuous scan experiment at 100 mM KCl. The obtained profile, shown in **Figure 8c (top panel)**, exhibited a broad spread signature, which is characteristic for bulk monomeric protein.

Sizing of the profile **(Figure 8c, bottom panel)** yielded a hydrodynamic radius of R_H = 4.29 ± 0.8 nm, which is in agreement with the size of monomeric GFP-tagged TDP-43. More importantly, in addition to the characteristic signature for monomeric protein in the continuous scan diffusion profile, we observed bright bursts that were narrowly distributed at the center of the channel on top of the diffusion profile, indicating the presence of clusters **(Figure 8c, top panel)**. To explore this further, we carried out smMDS step scan measurements. We performed high-resolution step scans with an interval of only 1 µm between steps within the central region of the diffusion profile where clusters would appear. Outside these regions we performed step scans with lower resolution **(Figure 8e, top panel)**. Clusters were clearly detectable as bursts above the mean bulk signal in the fluorescence time traces **(Figure 8d)**. These bursts were then counted to give us the number of clusters present at each position in the channel and binned in a histogram **(Figure 8e, bottom panel)**. Each peak in the histogram was then independently fit to a Gaussian distribution to obtain a mean diffusion distance that could be utilized to calculate D and, thus, R_H. Specifically, we extracted the half of the full-width half maximum (FWHM) of the Gaussians as the diffused distance at the four channel positions, corresponding to 0, 10, 26, and 55 seconds of travel within the channel. The diffusion distances x at each time point t **(Figure 8f)** were then fitted with a one-dimensional approximation of Fick's law

$$x \approx \sqrt{2Dt} \tag{1}$$

to extract D and, thus, R_H via the Stokes–Einstein relation. From the fit, using the average of three replicate measurements, we obtained an R_H of 121.5 ± 14.5 nm for TDP-43 nanoclusters **(Figure 8f)**. We note that the sizes of clusters are similar to FUS clusters previously observed,[91] thus indicating that TDP-43 forms similar pre-phase separation clusters as FUS.

The simultaneous, yet independent measurement of the sizes of nanoclusters and monomeric protein further allows estimating the number of monomer units per nanocluster. This is done through comparison of the volume ratios of monomeric TDP-43 and the clustered form. Assuming no restructuring of the protein within the nanocluster, a single cluster could contain as much as 20,000 proteins if the cluster is composed of pure protein. However, as condensates are liquid in nature and contain solvent molecules, typical volume fractions of proteins within condensate systems are on the order of ~10–35%;[95–97] hence, we expect the number of proteins per cluster to be in the range of 2,000–7,000. In addition to size measurements, the ability to directly count clusters in a digital manner also enables the quantification of cluster particle concentrations and volume fractions. From three repeat measurements, we detected an average number of clusters of N = 2281 ± 929. Using a previously established conversion strategy,[46] this corresponds to a flux of F = 72,606 ± 29,570 clusters per second or a cluster particle concentration of c = 7.24 ± 2.9 pM, corresponding to a total nanocluster volume fraction of $\phi = 3.16 \cdot 10^{-5}$. The concentration of TDP-43 nanoclusters detected here was therefore more than an order of magnitude higher

than previously determined for FUS nanoclusters formed under the same protein and salt concentrations,[91] suggesting a difference in the intermolar interactions that stabilize TDP-43 nanoclusters. Notably, TDP-43 is prone to aggregation and possesses a disordered domain capable of forming amyloid fibrils.[98] These characteristics may contribute to enhanced intermolecular interactions also in the clustered state, and potentially also explain the higher propensity of TDP-43 to form nanocluster assemblies.

Taken together, we have shown here that smMDS is able to capture and size pre-phase separation nanoclusters formed beyond the detection limit of conventional fluorescence microscopy. The discovery of TDP-43 nanoclusters and the understanding of the nature of such sub-diffraction assemblies is critical, in particular, for progressing our understanding of macroscopic phase separation phenomena. Our single-molecule sizing approach therefore provides insight into a largely unexplored area of protein assembly, taking advantage of the capability to elucidate properties of low abundance species present in a heterogeneous condensate system.

Conclusions

The physical characterization of the sizes, interactions, and assembly states of proteins is vital for a heightened understanding of the biological function, malfunction, and therapeutic intervention of proteins. Of particular interest are insights into the compositional heterogeneity of proteins and protein mixtures such as of protein oligomers, protein aggregation reactions or protein condensate systems. However, such characterization is challenging to achieve using conventional biophysical approaches as these methods are mostly reliant on ensemble readouts, which limits sensitivity and the information content that can be retrieved from such measurements. The smMDS approach developed herein addresses these challenges and takes advantage of the high sensitivity afforded by single-molecule detection and the physical features of diffusion in the microfluidic regime to enable digital sizing of proteins, protein assemblies, and heterogenous multicomponent protein systems directly in solution.

With different examples, ranging from single proteins to protein assemblies, we illustrated how the digital nature of the smMDS approach enables diffusional-sizing-based monitoring of protein hydrodynamic radii down to the femtomolar concentration range, thereby pushing the limits of sizing experiments by more than 5 orders of magnitude. smMDS further enables measurement of binding affinities of protein interactions at the single-molecule level and allows resolving high- and low-oligomeric states of proteins to gain insights into subpopulation distributions and oligomer equilibria. We further characterized the polydisperse nature of protein assemblies on the example of a protein aggregation reaction and applied smMDS to discover nanoclusters of a phase separating condensate system. These examples highlight the capability of the approach to elucidate properties of low abundance species present in heterogeneous biomolecular systems.

The implementation of the smMDS platform as demonstrated here is based on the integration of microfluidic diffusional sizing with in-situ scanning confocal microscopy. Future iterations of the platform could also incorporate, for example, multicolor single-molecule detection and FRET techniques,[99–101] as well as other, downstream microfluidic separation modalities,[102] to increase sensitivity, resolution, and information content even further. We also foresee that smMDS has the potential to be developed into a commercial benchtop instrument, similar to the one developed, for example, by Fluidic Analytics for conventional MDS experiments.[103] We note that the analytes tested in this work here were probed in aqueous buffer solutions. However, the smMDS approach should be adaptable to the detection of targets in biological and clinical samples. We further note that analytes measured here are in the 1–100 nm regime. However, smMDS is not limited to this size range and much larger proteins and protein assemblies can be quantified by adapting, for example, the flow rates or the chip design.

Taken together, the new capabilities of the smMDS platform augment the information content from sizing experiments beyond what is achievable and assayable with classical techniques. smMDS not only enables direct digital sizing of proteins and protein assemblies, but also provides quantitative information on protein interactions and heterogenous multicomponent protein systems. Given the key features of the technique, we anticipate that the smMDS approach will have a multitude of applications in quantifying the sizes and interactions of proteins and other biomolecules in various areas of biological and biomedical research, including the mechanistic and functional analysis of proteins, the molecular design of protein therapeutics, and the characterization of new nanomedicines and biomaterials.

Methods

Protein and sample preparation

Alexa 488 carboxylic acid was obtained as lyophilized powder from Thermo Fisher. Stock solutions at millimolar concentrations were prepared in dimethyl sulfoxide and further diluted in PBS buffer. Lysozyme, thyroglobulin, HSA, and RNase A were purchased from Sigma Aldrich as lyophilized powder in the highest purity available and suspended in 100 mM $NaHCO_3$ buffer (pH 8.2). Human Leukocyte antigen (HLA) A*03:01 was obtained through the NIH Tetramer Core Facility at Emory University (USA) and rebuffered into 100 mM $NaHCO_3$ buffer (pH 8.2). Human wildtype α-synuclein was recombinantly produced following a protocol detailed elsewhere and prepared in PBS.[104] TDP-43 was produced as a C-terminal EGFP-tagged protein variant in insect cells as previously described[88] and stored in 50 mM Tris-HCl (pH 7.4), 500 mM KCl, 5% (w/v) glycerol, 1 mM DTT. Concentrations were determined by UV/VIS spectroscopy.

Protein labeling

Protein solutions (lysozyme, thyroglobulin, HSA, RNase A, HLA, and α-synuclein) were incubated with *N*-hydroxysuccinimide (NHS)-ester-functionalized Alexa 488

(Thermo Fisher) at 5-molar excess of dye and incubated at room temperature for 1 h. Labeled protein was then separated from unbound dye by size-exclusion chromatography on a Superdex 200 increase 10/300 GL column (Cytiva) connected to an AKTA pure chromatography system (Cytiva) with PBS as elution buffer. For labeling of α-synuclein, the incubation was done at 4°C for 16 h and purification was performed on a Superdex 75 increase 10/300 GL column (Cytiva). Following separation, the labeled protein concentrations were determined using UV/VIS spectroscopy. The proteins were stored at –80°C until further use.

Generation of labeled α-synuclein oligomers and fibrils

A cysteine-containing variant (N122C) of α-synuclein was used for the preparation of labeled α-synuclein oligomers. The protein variant was produced as previously described[59] and labeled with maleimide-functionalized Alexa 488 dye, followed by purification using Sephadex G25 column (GE Healthcare). Oligomers were produced according to procedures detail elsewhere.[59] Briefly, labeled monomeric α-synuclein was lyophilized in deionized water and subsequently resuspended in PBS (pH 7.4) at a final concentration of ~6 mg mL^{-1}. The resulting solution was passed through a 0.22 µm filter (Millipore) before incubation at 37°C for 16 h under quiescent conditions. Small amounts of fibrillar species were removed by ultracentrifugation. Excess monomeric protein was then removed by multiple filtration steps using 100-kDa cut-off membranes. The final oligomer concentration was determined by UV/VIS spectroscopy.

Recombinant human wildtype α-synuclein was used for the generation of α-synuclein fibrils. Fibrils were prepared from a mixture of Alexa 488 labeled and unlabeled α-synuclein protein, as previously described.[30] Briefly, a mixture containing 10% labeled and 90% unlabeled α-synuclein monomer was shaken at 37 °C and 200 rpm for 4 days to generate 1st generation fibrils. Then, the seeds of 1st generation fibrils were recovered by centrifugation and incubated with 10% of labeled and 90% unlabeled monomer at 37 °C and 200 rpm for 3 days to generate 2nd generation fibrils. After centrifugation and recovery, 2nd generation fibrils were sonicated (10% power, 30% cycles for 90 s) using a Sonopuls HD 2070 ultrasonic homogenizer (Bandelin) and stored at room temperature until further use. Fibril concentration was determined by UV/VIS spectroscopy.

smMDS platform

The approach described here integrates microchip-based diffusional sizing with confocal fluorescence detection. Schematics of the microfluidic device, the optical setup, and their integration are shown in **Figure 1a**. Briefly, the microfluidic chip design is based on previously reported device designs for diffusional sizing.[25,30,105] The device has two inlets, one for the injection of the sample and one for the injection of co-flowing buffer solution. Channels of 25 µm in height and 50 µm in width, respectively, direct the sample and the buffer solutions to an entry nozzle. At the entry nozzle, the sample and the buffer stream merge into an observation channel of 25 µm in height and

225 µm in width, in which diffusion is monitored. Notably, the channel geometry at the nozzle point is designed such that sample and buffer solution are drawn through the chip in a ~1:8 volume ratio. The observation channel is folded multiple times and is approximately 90'000 µm long and terminates at a waste outlet where negative pressure is applied by a syringe pump. Scanning markers are integrated on the chip adjacent to the observation channel for defining start and end points of the scan trajectory. Details on the fabrication of the device by standard soft-lithography and molding techniques are given below.

The optical unit of the smMDS platform is based on fluorescence confocal microscopy and optimized for microfluidic experiments. The microscope is built around a 'rapid automated modular microscope' (RAMM) frame (Applied Scientific Instrumentation (ASI)) and is equipped with a motorized x,y,z-scanning stage (PZ-2000FT, ASI), onto which the diffusional sizing chip is mounted. For controlling the exact sample placement along the optical axis of the microscope, the stage is equipped with a z-piezo. To excite the sample in the device, the beam of a 488-nm wavelength laser (Cobolt 06-MLD, 200 mW diode laser, Cobolt) is passed through a single-mode optical fiber (P3-488PM-FC-1, Thorlabs) and collimated at the exit of the fiber by an achromatic collimator (60FC-L-4-M100S-26, Schäfter + Kirchhoff) to form a beam with a Gaussian profile. The beam is then directed into the microscope, reflected by a dichroic beamsplitter (Di03-R488/561, Semrock), and subsequently focused to a concentric diffraction-limited spot in the microfluidic channel through a 60x-magnification water-immersion objective (CFI Plan Apochromat WI 60x, NA 1.2, Nikon). The emitted light from the sample is collected via the same objective, passed through the dichroic beam splitter, and focused by achromatic lenses through a 30-µm pinhole (Thorlabs) to remove any out-of-focus light. The emitted photons are filtered through a band-pass filter (FF01-520/35-25, Semrock) and then focused onto a single-photon counting avalanche photodiode (APD, SPCM-14, PerkinElmer Optoelectronics), which is connected to a TimeHarp260 time-correlated single photon counting unit (PicoQuant).

Fabrication of microfluidic devices

The microfluidic device for smMDS was fabricated in poly(dimethylsiloxane) (PDMS) through a single, standard soft-photolithography step. In brief, the device was designed in AutoCAD software (Autodesk) and printed on acetate transparencies (Micro Lithography Services). The master replica for fabricating the device was prepared by soft-lithography methods.[106] Accordingly, SU-8 3025 (Kayaku) was deposited onto a polished silicon wafer and then spun to a height of approximately 25 µm. The acetate mask was placed onto the coated wafer and exposed with UV light using a custom-built LED-based apparatus.[107] Following UV exposure, the master was developed in propylene glycol methyl ether acetate (PGMEA, Sigma Aldrich). The exact height was measured by a Dektak profilometer (Bruker). The mold was then used to generate a patterned PDMS chip. For this purpose, PDMS (Sylgard 184, Dow Corning) was mixed with curing agent (Sylgard 184, Dow Corning) at a ratio of 10:1

(w/w) and subsequently degassed and baked for approximately 1 h at 65°C. This PDMS cast was removed from the mold, and access holes for the inlet and outlet connectors were punched with biopsy punches. The devices were bonded onto a thin glass coverslip (no. 1.5, Menzel) after both the PDMS and the coverslip glass surface had been activated by oxygen plasma (Diener electronic, 40 % power for 30 s). Before injecting the buffer into the channels, the chips were rendered more hydrophilic through an additional plasma oxidation step (Diener electronic, 80% power for 500 s).[108]

Experimental procedures

All experiments were performed at room temperature. Buffer was PBS (pH 7.4) in all experiments, except in nanocluster experiments, where TRIS buffer was used (50 mM TRIS-HCl, pH 7.4). Buffers were supplemented with 0.01% Tween 20 (Thermo Fisher) to prevent adhesion of molecules to chip surfaces. The PDMS–glass device was secured to the motorized, programmable microscope stage. Co-flowing buffer and sample were entered into the chip through gel-loading tips inserted into the appropriate inlet orifices and drawn through the chip by applying negative pressure with a syringe (Hamilton) and syringe pump (Cetoni, neMESYS) connected to the outlet. Flow rates were 100 µL/h in all experiments, except in nanocluster experiments, where the flow rates were 150 µL/h for continuous scan experiments and 60 µL/h for step scan experiments. Flow was allowed to equilibrate over six minutes before data acquisition. Diffusion profiles were obtained by translocating the confocal volume either in a continuous or stepwise manner through the four innermost channels of the microfluidic sizing chip using a custom-written Python script that simultaneously controlled the stage movement and the data acquisition at the mid-height of the device (i.e., ~12.5 µm above the surface of the glass coverslip). Continuous scans were performed at 20–100 µm/s. Step scans were done in 200–400 steps for a duration of 1–60 s at each position. The scanning markers were used to define the x,y,z-coordinates of the start and end positions of the scan trajectory. Each experiment was performed in a freshly fabricated PDMS device. The laser power at the back aperture of the objective was adjusted to 370 µW in all experiments, except for experiments on fibrils, where laser powers of 100 µW were used, and for experiments on nanoclusters, where laser powers of 6 µW were used. Photon recordings were done in T2 mode and the arrival times of photons were measured with respect to the overall measurement start with 16-picosecond resolution.

Data analysis

Data analysis and plotting was done in Python. In continuous scan experiments, photon recordings were binned in 1-ms intervals to obtain intensity readouts, from which diffusion profiles were generated by plotting the obtained fluorescence intensities as a function of chip position. In step scan experiments, diffusion profiles were created by extracting single-molecule events from the recorded time trace at each position using a burst-search algorithm, and plotting the obtained the number of counted molecules as a function of chip position. The burst-search algorithm identifies

single molecules from the photon time trace by applying a combined maximum inter-photon time (IPT_{max}) and minimum total number of photons (N_{min}) threshold. IPT_{max} and N_{min} were in the range of 0.005–0.02 ms and 5–20 number of photons, respectively, for all experiments performed under single-molecule conditions, unless otherwise stated. In addition, a Lee filter of 2–4 was applied that smoothens regions of constant signal while keeping those with rapid parameter changes unaffected (such as the edges of the bursts).

To extract size information, the obtained diffusion profiles, from both continuous and step-scan experiments, were analyzed with a custom-written analysis software, freely available on Github (https://github.com/impact27/diffusion_device). This script fits the obtained diffusion profiles with simulated diffusion profiles from numerical model simulations solving the diffusion–advection equations for mass transport under flow. Hereby, a least-squares error algorithm is used to find simulated profiles with the lowest residuals to determine D and recover R_H via the Stokes–Einstein relationship.

Conflict of interest

The authors declare no competing interests.

Author contributions

G.K. and T.P.J.K. conceptualized the study, G.K. devised the methodology, G.K., M.M.S., T.J.W., J.F., E.A.A., G.S., M.A.C., H.A., L.C. performed the investigation, G.K., R.P.B.J., M.M.S., T.J.W., Q.A.E.P. performed the analysis, R.P.B.J. and Q.A.E.P. contributed relevant software, W.E.A., K.L.S., T.W.H., T.M.F., V.K., S.A., F.-U. H., S.F.L. provided materials and expertise. G.K. wrote the original draft and all other authors reviewed and edited it.

Acknowledgements

The research leading to these results has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 Framework Programme through the Marie Skłodowska-Curie grant MicroSPARK (agreement no. 841466; G.K.), the Herchel Smith Fund of the University of Cambridge (G.K.), the Wolfson College Junior Research Fellowship (G.K.), the European Innovation Council (EIC) under the European Union's Horizon 2020 Future and Emerging Technologies Open (FET-OPEN) programme through the grant NANOPHLOW (agreement no. 766972; T.P.J.K.), the European Union's Horizon 2020 research and innovation programme through the grant NANOTRANS (agreement no. 674979; Q.A.E.P., T.P.J.K.), and the Newman Foundation (T.P.J.K.). T.J.W acknowledges funding from the Harding Distinguished Postgraduate Scholar Programme, administered by the Cambridge Commonwealth, European & International Trust at the University of Cambridge. K.L.S. acknowledges funding from the Schmidt Science Fellowship in partnership with the Rhodes Trust and the St. John's College Junior Research Fellowship. V.K. acknowledges funding from an NIHR Fellowship (PDF-2016-09-065) and as a Paul I. Terasaki Scholar. This work was also supported by a grant from the

Engineering and Physical Sciences Research Council (EPSRC) (grant no. EP/L015889/1; R.P.B.J.). The authors would also like to thank the EPSRC Centre for Interdisciplinary Ph.D. Training and Research in Nanoscience and Nanotechnology (NanoDTC).

References

1. Goodsell, D. S. *The Machinery of Life*. *The Machinery of Life* (Springer New York, 2009). doi:10.1007/978-0-387-84925-6.

2. Neet, K. E. & Lee, J. C. Biophysical characterization of proteins in the post-genomic era of proteomics. *Mol Cell Proteomics* **1**, 415–420 (2002).

3. Knowles, T. P. J., Vendruscolo, M. & Dobson, C. M. The amyloid state and its association with protein misfolding diseases. *Nat Rev Mol Cell Biol* **15**, 384–396 (2014).

4. Eliezer, D. Biophysical characterization of intrinsically disordered proteins. *Curr Opin Struct Biol* **19**, 23–30 (2009).

5. Bemporad, F. & Chiti, F. Protein Misfolded Oligomers: Experimental Approaches, Mechanism of Formation, and Structure-Toxicity Relationships. *Chem Biol* **19**, 315–327 (2012).

6. Ha, J., Park, H., Park, J. & Park, S. B. Recent advances in identifying protein targets in drug discovery. *Cell Chem Biol* **28**, 394–423 (2021).

7. Rapposelli, S., Gaudio, E., Bertozzi, F. & Gul, S. Editorial: Protein–Protein Interactions: Drug Discovery for the Future. *Front Chem* **9**, 1049 (2021).

8. Shen, Y. *et al.* From Protein Building Blocks to Functional Materials. *ACS Nano* **15**, 5819–5837 (2021).

9. Houde, D. J. & Berkowitz, S. A. *Biophysical Characterization of Proteins in Developing Biopharmaceuticals*. *Biophysical Characterization of Proteins in Developing Biopharmaceuticals* (Elsevier, 2020). doi:10.1016/C2017-0-03008-2.

10. Renaud, J. P. *et al.* Biophysics in drug discovery: impact, challenges and opportunities. *Nature Reviews Drug Discovery 2016 15:10* **15**, 679–698 (2016).

11. Hamley, I. W. Protein Assemblies: Nature-Inspired and Designed Nanostructures. *Biomacromolecules* **20**, 1829–1848 (2019).

12. Modena, M. M., Rühle, B., Burg, T. P. & Wuttke, S. Nanoparticle Characterization: What to Measure? *Advanced Materials* **31**, 1901556 (2019).

13. Lin, P. C., Lin, S., Wang, P. C. & Sridhar, R. Techniques for physicochemical characterization of nanomaterials. *Biotechnol Adv* **32**, 711 (2014).

14. Walport, L. J., Low, J. K. K., Matthews, J. M. & MacKay, J. P. The characterization of protein interactions – what, how and how much? *Chem Soc Rev* **50**, 12292–12307 (2021).

15. Tamara, S., Den Boer, M. A. & Heck, A. J. R. High-Resolution Native Mass Spectrometry. *Chem Rev* **122**, 7269–7326 (2022).

16. Barth, H. G., Jackson, Christian. & Boyes, B. E. Size Exclusion Chromatography. *Anal Chem* **66**, 595–620 (1994).

17. Nguyen, H. H., Park, J., Kang, S. & Kim, M. Surface plasmon resonance: a versatile technique for biosensor applications. *Sensors (Basel)* **15**, 10481–10510 (2015).

18. Fekete, S., Guillarme, D., Sandra, P. & Sandra, K. Chromatographic, Electrophoretic, and Mass Spectrometric Methods for the Analytical Characterization of Protein Biopharmaceuticals. *Anal Chem* **88**, 480–507 (2016).

19. Velázquez Campoy, A. & Freire, E. ITC in the post-genomic era...? Priceless. *Biophys Chem* **115**, 115–24 (2005).

20. Cole, J. L., Lary, J. W., P. Moody, T. & Laue, T. M. Analytical Ultracentrifugation: Sedimentation Velocity and Sedimentation Equilibrium. *Methods Cell Biol* **84**, 143 (2008).

21. Stetefeld, J., McKenna, S. A. & Patel, T. R. Dynamic light scattering: a practical guide and applications in biomedical sciences. *Biophys Rev* **8**, 409 (2016).

22. Arter, W. E., Levin, A., Krainer, G. & Knowles, T. P. J. Microfluidic approaches for the analysis of protein–protein interactions in solution. *Biophys Rev* 1–11 (2020) doi:10.1007/s12551-020-00679-4.

23. Charmet, J., Arosio, P. & Knowles, T. P. J. Microfluidics for Protein Biophysics. *J Mol Biol* **430**, 565–580 (2018).

24. Otzen, D. E., Buell, A. K. & Jensen, H. Microfluidics and the quantification of biomolecular interactions. *Curr Opin Struct Biol* **70**, 8–15 (2021).

25. Arosio, P. *et al.* Microfluidic diffusion analysis of the sizes and interactions of proteins under native solution conditions. *ACS Nano* **10**, 333–341 (2016).

26. Zhang, Y. *et al.* On-chip measurements of protein unfolding from direct observations of micron-scale diffusion. *Chem Sci* **9**, 3503–3507 (2018).

27. Challa, P. K. *et al.* Real-Time Intrinsic Fluorescence Visualization and Sizing of Proteins and Protein Complexes in Microfluidic Devices. *Anal Chem* **90**, 3849–3855 (2018).

28. Wright, M. A. *et al.* Cooperative Assembly of Hsp70 Subdomain Clusters. *Biochemistry* **57**, 3641–3649 (2018).

29. Fiedler, S. *et al.* Antibody Affinity Governs the Inhibition of SARS-CoV-2 Spike/ACE2 Binding in Patient Serum. *ACS Infect Dis* **7**, 2362–2369 (2021).

30. Schneider, M. M. *et al.* The Hsc70 disaggregation machinery removes monomer units directly from α-synuclein fibril ends. *Nature Communications 2021 12:1* **12**, 1–11 (2021).

31. Linse, S. *et al.* Kinetic fingerprints differentiate the mechanisms of action of anti-Aβ antibodies. *Nat Struct Mol Biol* **27**, 1125–1133 (2020).

32. Baron, J. *et al.* FULL-MDS: Fluorescent Universal Lipid Labeling for Microfluidic Diffusional Sizing. *Anal Chem* **95**, 587–593 (2023).

33. Fiedler, S. *et al.* Serological fingerprints link antiviral activity of therapeutic antibodies to affinity and concentration. *Scientific Reports 2022 12:1* **12**, 1–10 (2022).

34. Scheidt, T. *et al.* Secondary nucleation and elongation occur at different sites on Alzheimer's amyloid-b aggregates. *Sci Adv* **5**, (2019).

35. Yates, E. V. *et al.* Latent analysis of unmodified biomolecules and their complexes in solution with attomole detection sensitivity. *Nat Chem* **7**, 802–809 (2015).

36. Moser, M. R. & Baker, C. A. Taylor dispersion analysis in fused silica capillaries: a tutorial review. *Analytical Methods* **13**, 2357–2373 (2021).

37. Poulsen, N. N. *et al.* Flow induced dispersion analysis rapidly quantifies proteins in human plasma samples. *Analyst* **140**, 4365–4369 (2015).

38. Walt, D. R. Optical methods for single molecule detection and analysis. *Anal Chem* **85**, 1258–1263 (2013).

39. Deniz, A. A. Deciphering Complexity in Molecular Biophysics with Single-Molecule Resolution. *J Mol Biol* **428**, 301–307 (2016).

40. Weiss, S. Fluorescence spectroscopy of single biomolecules. *Science* **283**, 1676–83 (1999).

41. Kapanidis, A. N. & Strick, T. Biology, one molecule at a time. *Trends Biochem Sci* **34**, 234–243 (2009).

42. Horrocks, M. H. *et al.* Single molecule fluorescence under conditions of fast flow. *Anal Chem* **84**, 179–85 (2012).

43. Zijlstra, N. *et al.* Rapid Microfluidic Dilution for Single-Molecule Spectroscopy of Low-Affinity Biomolecular Complexes. *Angewandte Chemie International Edition* **56**, 7126–7129 (2017).

44. Tyagi, S. *et al.* Continuous throughput and long-term observation of single-molecule FRET without immobilization. *Nat Methods* **11**, 297–300 (2014).

45. Kim, S. *et al.* High-throughput single-molecule optofluidic analysis. *Nat Methods* **8**, 242–5 (2011).

46. Krainer, G. *et al.* Direct digital sensing of protein biomarkers in solution. *Nature Communications 2023 14:1* **14**, 1–21 (2023).

47. Haustein, E. & Schwille, P. Fluorescence Correlation Spectroscopy: Novel Variations of an Established Technique. *Annu Rev Biophys Biomol Struct* **36**, 151–169 (2007).

48. Yu, L. *et al.* A Comprehensive Review of Fluorescence Correlation Spectroscopy. *Front Phys* **9**, 644450 (2021).

49. Armstrong, J. K., Wenby, R. B., Meiselman, H. J. & Fisher, T. C. The hydrodynamic radii of macromolecules and their effect on red blood cell aggregation. *Biophys J* **87**, 4259–4270 (2004).

50. Jiang, X. *et al.* Quantitative analysis of the protein corona on FePt nanoparticles formed by transferrin binding. *J R Soc Interface* **7 Suppl 1**, (2010).

51. Akbarzadehlaleh, P., Mirzaei, M., Mashahdi-Keshtiban, M., Shamsasenjan, K. & Heydari, H. PEGylated Human Serum Albumin: Review of PEGylation, Purification and Characterization Methods. *Adv Pharm Bull* **6**, 309–317 (2016).

52. Dynamic Biosensors. List of protein hydrodynamic diameters DH. https://www.dynamic-biosensors.com/project/list-of (2022).

53. Vendelin, M. & Birkedal, R. Anisotropic diffusion of fluorescently labeled ATP in rat cardiomyocytes determined by raster image correlation spectroscopy. *Am J Physiol Cell Physiol* **295**, (2008).

54. Study of Protein Hydrodynamics with Light Scattering. https://www.brookhaveninstruments.com/study-of-protein-hydrodynamics-with-light-scattering-size-and-charge-of-lysozyme/.

55. Size Measurement of the Proteins RNAse A and Proteinase K. https://www.brookhaveninstruments.com/size-measurement-of-the-proteins-rnase-a-and-proteinase-k-and-the-effects-of-protein-binding/.

56. Falke, M. *et al.* α-Synuclein-derived lipoparticles in the study of α-Synuclein amyloid fibril formation. *Chem Phys Lipids* **220**, 57–65 (2019).

57. Schneider, M. M. *et al.* Microfluidic antibody affinity profiling of alloantibody-HLA interactions in human serum. *Biosens Bioelectron* **228**, 115196 (2023).

58. Schneider, S. *Determination of Protein Molecular Weight and Size Using the Agilent 1260 Infinity Multi-Detector Bio-SEC Solution with Advanced Light Scattering Detection*. *Application Note* (2014).

59. Arter, W. E. *et al.* Rapid Structural, Kinetic, and Immunochemical Analysis of Alpha-Synuclein Oligomers in Solution. *Nano Lett* **20**, 8163–8169 (2020).

60. Lakayan, D., Haselberg, R., Gahoual, R., Somsen, G. W. & Kool, J. Affinity profiling of monoclonal antibody and antibody-drug-conjugate preparations by coupled liquid chromatography-surface plasmon resonance biosensing. *Anal Bioanal Chem* **410**, 7837–7848 (2018).

61. Schneider, M. M. *et al.* Microfluidic characterisation reveals broad range of SARS-CoV-2 antibody affinity in human plasma. *Life Sci Alliance* **5**, e202101270 (2022).

62. Biddison, W. E. *et al.* Structural studies of an HLA-A03 alloantigenic epitope defined by a human hybridoma antibody. *Immunogenetics 1989 30:1* **30**, 54–57 (1989).

63. Ivanyi, D. & van de Meugheuvel, W. A monomorphic HLA-specific monoclonal antibody, W6/32, reacts with the H-2Db molecule of normal mouse lymphocytes. *Immunogenetics 1984 20:6* **20**, 699–703 (1984).

64. Levy, E. D. & Teichmann, S. Structural, evolutionary, and assembly principles of protein oligomerization. *Prog Mol Biol Transl Sci* **117**, 25–51 (2013).

65. Ali, M. H. & Imperiali, B. Protein oligomerization: how and why. *Bioorg Med Chem* **13**, 5013–5020 (2005).

66. Gell, D. A., Grant, R. P. & MacKay, J. P. The detection and quantitation of protein oligomerization. *Adv Exp Med Biol* **747**, 19–41 (2012).

67. Kulenkampff, K., Wolf Perez, A. M., Sormanni, P., Habchi, J. & Vendruscolo, M. Quantifying misfolded protein oligomers as drug targets and biomarkers in Alzheimer and Parkinson diseases. *Nature Reviews Chemistry 2021 5:4* **5**, 277–294 (2021).

68. Kumari, N. & Yadav, S. Modulation of protein oligomerization: An overview. *Prog Biophys Mol Biol* **149**, 99–113 (2019).

69. Griffin, M. D. W. & Gerrard, J. A. The relationship between oligomeric state and protein function. *Adv Exp Med Biol* **747**, 74–90 (2012).

70. Young, G. *et al.* Quantitative mass imaging of single biological macromolecules. *Science (1979)* **360**, 423–427 (2018).

71. Puchalla, J., Krantz, K., Austin, R. & Rye, H. Burst analysis spectroscopy: A versatile single-particle approach for studying distributions of protein

aggregates and fluorescent assemblies. *Proc Natl Acad Sci U S A* **105**, 14400–14405 (2008).

72. Gambin, Y. *et al.* Confocal Spectroscopy to Study Dimerization, Oligomerization and Aggregation of Proteins: A Practical Guide. *International Journal of Molecular Sciences 2016, Vol. 17, Page 655* **17**, 655 (2016).

73. Du, X. Y., Xie, X. X. & Liu, R. T. The Role of α-Synuclein Oligomers in Parkinson's Disease. *International Journal of Molecular Sciences 2020, Vol. 21, Page 8645* **21**, 8645 (2020).

74. Musteikytė, G. *et al.* Interactions of α-synuclein oligomers with lipid membranes. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1863**, 183536 (2020).

75. Alam, P., Bousset, L., Melki, R. & Otzen, D. E. α-synuclein oligomers and fibrils: a spectrum of species, a spectrum of toxicities. *J Neurochem* **150**, 522–534 (2019).

76. Chen, S. W. *et al.* Structural characterization of toxic oligomers that are kinetically trapped during α-synuclein fibril formation. *Proc Natl Acad Sci U S A* **112**, E1994–E2003 (2015).

77. Lorenzen, N. *et al.* The role of stable α-synuclein oligomers in the molecular events underlying amyloid formation. *J Am Chem Soc* **136**, 3859–3868 (2014).

78. Fusco, G. *et al.* Structural basis of membrane disruption and cellular toxicity by α-synuclein oligomers. *Science (1979)* **358**, 1440–1443 (2017).

79. Lee, J. E. *et al.* Mapping surface hydrophobicity of α-synuclein oligomers at the nanoscale. *Nano Lett* **18**, 7494–7501 (2018).

80. Staats, R. *et al.* Screening of small molecules using the inhibition of oligomer formation in α-synuclein aggregation as a selection parameter. *Communications Chemistry 2020 3:1* **3**, 1–9 (2020).

81. Ross, C. A. & Poirier, M. A. Protein aggregation and neurodegenerative disease. *Nat Med* **10**, S10–S17 (2004).

82. Lashuel, H. A. Do Lewy bodies contain alpha-synuclein fibrils? and Does it matter? A brief history and critical analysis of recent reports. *Neurobiol Dis* **141**, 104876 (2020).

83. Meisl, G. *et al.* Differences in nucleation behavior underlie the contrasting aggregation kinetics of the Aβ40 and Aβ42 peptides. *Proceedings of the National Academy of Sciences* **111**, 9384–9389 (2014).

84. Shin, Y. & Brangwynne, C. P. Liquid phase condensation in cell physiology and disease. *Science (1979)* **357**, (2017).

85. Alberti, S. & Dormann, D. Liquid–Liquid Phase Separation in Disease. *Annu Rev Genet* **53**, 171–194 (2019).

86. Brangwynne, C. P., Tompa, P. & Pappu, R. V. Polymer physics of intracellular phase transitions. *Nat Phys* **11**, 899–904 (2015).

87. Boeynaems, S. *et al.* Protein Phase Separation: A New Phase in Cell Biology. *Trends Cell Biol* **28**, 420–435 (2018).

88. Krainer, G. *et al.* Reentrant liquid condensate phase of proteins is stabilized by hydrophobic and non-ionic interactions. *Nature Communications 2021 12:1* **12**, 1–14 (2021).

89. Maharana, S. *et al.* RNA buffers the phase separation behavior of prion-like RNA binding proteins. *Science* **360**, 918–921 (2018).

90. Patel, A. *et al.* A Liquid-to-Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation. *Cell* **162**, 1066–1077 (2015).

91. Kar, M. *et al.* Phase-separating RNA-binding proteins form heterogeneous distributions of clusters in subsaturated solutions. *Proc Natl Acad Sci U S A* **119**, e2202222119 (2022).

92. Wen, J. *et al.* Conformational Expansion of Tau in Condensates Promotes Irreversible Aggregation. *J Am Chem Soc* **143**, 13056–13064 (2021).

93. Li, P. *et al.* Phase transitions in the assembly of multivalent signalling proteins. *Nature* **483**, 336–340 (2012).

94. Seim, I. *et al.* Dilute phase oligomerization can oppose phase separation and modulate material properties of a ribonucleoprotein condensate. *Proc Natl Acad Sci U S A* **119**, (2022).

95. Burke, K. A., Janke, A. M., Rhine, C. L. & Fawzi, N. L. Residue-by-Residue View of In Vitro FUS Granules that Bind the C-Terminal Domain of RNA Polymerase II. *Mol Cell* **60**, 231–241 (2015).

96. Murthy, A. C. *et al.* Molecular interactions underlying liquid−liquid phase separation of the FUS low-complexity domain. *Nature Structural & Molecular Biology 2019 26:7* **26**, 637–648 (2019).

97. Ahlers, J. *et al.* The key role of solvent in condensation: Mapping water in liquid-liquid phase-separated FUS. *Biophys J* **120**, 1266–1275 (2021).

98. Li, Q., Babinchak, W. M. & Surewicz, W. K. Cryo-EM structure of amyloid fibrils formed by the entire low complexity domain of TDP-43. *Nature Communications 2021 12:1* **12**, 1–8 (2021).

99. Hohlbein, J., Craggs, T. D. & Cordes, T. Alternating-laser excitation: single-molecule FRET and beyond. *Chem. Soc. Rev.* **43**, 1156–1171 (2014).

100. Schuler, B. Single-molecule FRET of protein structure and dynamics - a primer. *J Nanobiotechnology* **11**, S2 (2013).

101. Orte, A., Clarke, R. & Klenerman, D. Single-molecule two-colour coincidence detection to probe biomolecular associations. *Biochem Soc Trans* **38**, 914–918 (2010).

102. Saar, K. L. *et al.* Rapid two-dimensional characterisation of proteins in solution. *Microsyst Nanoeng* **5**, 1–10 (2019).

103. Fluidic Analytics. https://www.fluidic.com/.

104. Hoyer, W. *et al.* Dependence of α-synuclein aggregate morphology on solution conditions. *J Mol Biol* **322**, 383–393 (2002).

105. Herling, T. W. *et al.* Non-specificity fingerprints for clinical stage antibodies in solution. *bioRxiv* 2023.02.13.528263 (2023) doi:10.1101/2023.02.13.528263.

106. Duffy, D. C., McDonald, J. C., Schueller, O. J. A. & Whitesides, G. M. Rapid prototyping of microfluidic systems in poly(dimethylsiloxane). *Anal Chem* **70**, 4974–4984 (1998).

107. Challa, P. K., Kartanas, T., Charmet, J. & Knowles, T. P. J. Microfluidic devices fabricated using fast wafer-scale LED-lithography patterning. *Biomicrofluidics* **11**, 014113 (2017).

108. Tan, S. H., Nguyen, N.-T., Chua, Y. C. & Kang, T. G. Oxygen plasma treatment for reducing hydrophobicity of a sealed polydimethylsiloxane microchannel. *Biomicrofluidics* **4**, 032204 (2010).