ACS Photonics

pubs.acs.org/journal/apchd5

Article

A Quantitative Description for Optical Mass Measurement of Single Biomolecules

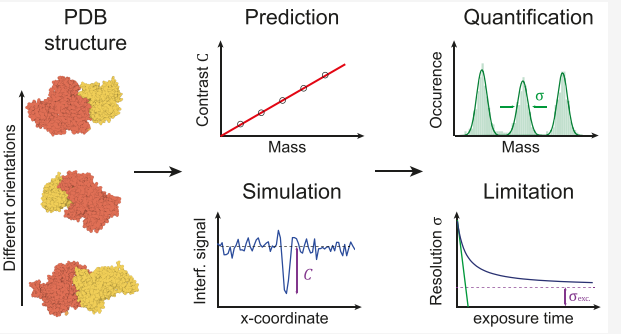
Jan Becker,* Jack S. Peters, Ivor Crooks, Seham Helmi, Marie Synakewicz, Benjamin Schuler, and Philipp Kukura*

Cite This: https://doi.org/10.1021/acsphotonics.3c00422

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Label-free detection of single biomolecules in solution has been achieved using a variety of experimental approaches over the past decade. Yet, our understanding of the magnitude of the optical contrast and its relationship with the underlying atomic structure as well as the achievable measurement sensitivity and precision remain poorly defined. Here, we use a Fourier optics approach combined with an atomic structure-based molecular polarizability model to simulate mass photometry experiments from first principles. We find excellent agreement between several key experimentally determined parameters such as optical contrast-to-mass conversion, achievable mass accuracy, and molecular shape and orientation dependence. This allows us to determine detection sensitivity and measurement precision mostly independent of the optical detection approach chosen, resulting in a general framework for light-based single-molecule detection and quantification.



KEYWORDS: *mass photometry, polarizability, single molecule, label free, mass measurement*

1. INTRODUCTION

Recent developments in ultrasensitive light microscopy[1,2] have enabled the quantification of biomolecular mass, charge, and size at the single-molecule level and in solution.[3−8] In mass photometry (MP), light scattered from a protein when it binds to or moves along a glass coverslip in solution is detected together with partially reflected light from the glass−water interface (Figure 1a). MP has demonstrated both high mass accuracy and precision on the order of a few percent of the object mass, enabled by high-measurement precision at the single-molecule level.[3] This has been achieved through selective attenuation of the reflected light using a mask in the back-focal-plane (BFP) of the optical system,[9] coupled with averaging of detected photoelectrons by the imaging camera and post-processing of the raw images, enabling the detection and resolution of oligomeric states and protein complexes.[10] As a result, MP can be used to quantify interaction affinities and kinetics,[11] molecular organization,[12] and for studies of biomolecular dynamics.[13]

The molecular mass, m, for unknown samples is inferred from the optical contrast of the molecule under investigation and an empirical scaling between the contrast and a species of known molecular mass. This relationship can be approximated by estimating the excess polarizability α using the refractive index of proteins n_p and assuming a spherical shape.[14] In fact, a number of approaches have been reported recently to calculate the optical signal in interferometric scattering microscopy, which can be combined with such simplified models of biomolecules to predict images and expected optical contrast.[15,16] Nevertheless, all models to date did not consider the atomic nature of biomolecules, making any attempts to compare experimental and theoretical results largely qualitative and unable to predict effects of molecular shape or orientation, for example. Given that the refractive index of a single protein is poorly defined, we thus lack a molecular-level description of light-based mass measurement and what properties define the limits and opportunities in measurement sensitivity, precision, and accuracy.

We thus set out to develop an approach capable of simulating images of individual proteins on a microscope cover glass in an MP instrument coupled with an explicit atomic description of molecular polarizability and thereby explore key aspects such as (1) to which degree optical contrasts reported to date experimentally match those predicted by theory. (2) How the measured signal depends on the molecular shape and orientation, thereby informing on the ultimately achievable mass accuracy and resolution. (3)

Received: March 28, 2023

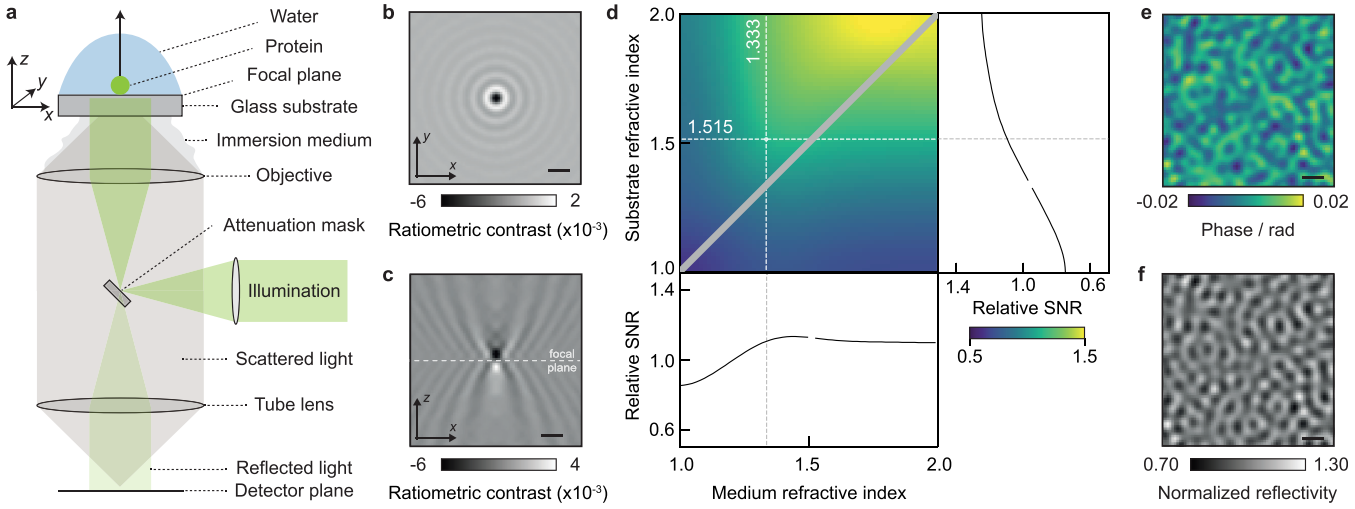


Figure 1. Fundamentals of image formation in a contrast-enhanced back reflection geometry. (a) Schematic of the simulated, widefield, MP setup, including an attenuation mask in the BFP of the objective lens, which selectively reduces light reflected from the glass coverslip. (b,c) Simulated ratiometric contrast for a single 24-mer of the small heat shock protein Hsp16.5 (m = 396 kDa) at an atomically flat glass–water interface. (d) Relative SNR when varying the refractive indices of the substrate and buffer medium. The diagonal indicates refractive index matching, where no reference field is available due to a lack of reflection. (e) Simulated phase retardation map at λ = 445 nm arising from nanoscopic roughness of glass coverslips on the order of ∼2 nm height variations over ∼100 nm (lateral) length scales. (f) Resulting speckle-like image using a 0.1% transmission mask, normalized to the expected reflectivity from a flat glass–water interface (see Section S6). Scalebars = 0.5 μm.

The current and likely future limits on measurement sensitivity and resolution for light-based single-molecule characterization for MP and beyond.

2. RESULTS

Our model simulates an experimental setup based on plane wave illumination, a (simplified) single lens high numerical aperture (NA) objective for light delivery and collection that is refractive index matched to the sample interface (here, glass; n_g), and a protein embedded in a medium of refractive index n_m (e.g., water). A spatial mask is used to selectively attenuate light reflected from the coverslip substrate, which increases the optical contrast, simplifies the accurate determination of the focal position for maximum contrast, and allows for higher illumination power, given limited camera full well capacities (Figure 1a).[11] We mathematically model the influence of the imaging system as the following convolution operation, defined as $f(x) \otimes g(x) = \int_{-\infty}^{+\infty} dx' f(x') \cdot g(x - x')$, with an amplitude point-spread-function (APSF)[17] $\vec{h}$

$$I = |[\vec{E}_{ref} + \vec{E}_{sca}] \otimes \vec{h}|^2 \quad (1)$$

yielding the detectable intensity I, with $\vec{E}_{ref}$ and $\vec{E}_{sca}$ being the reflected and scattered electric fields, respectively. These are directly linked to the illumination $\vec{E}_{illu}$ through

$$\vec{E}_{ref} \propto r \cdot \vec{E}_{illu} \quad (2)$$

$$\vec{E}_{sca} \propto t_2 \cdot s \cdot t_1 \cdot \vec{E}_{illu} \quad (3)$$

where r and t are the Fresnel coefficients for reflection and transmission,[18] and the subscripts indicate transmission of $\vec{E}_{illu}$ from glass → water (t_1) or that of the scattered field $\vec{E}_{sca}$ from water → glass (t_2). Note that the influence of the attenuation mask is realized by propagating both, $\vec{E}_{ref}$ and $\vec{E}_{sca}$, into the BFP of the objective lens, where they are being multiplied by a circular mask (corresponding to an effective NA, of 0.58) with a given transmission strength $|\tau|^2$ (1%; unless otherwise stated), motivated by experimental parameters.[9] The scattering coefficient s scales with the polarizability α of the protein, which in the Rayleigh regime[14] can be approximated to be proportional to the particle volume

$$s \propto \alpha \propto V \cdot \frac{n_p^2 - n_m^2}{n_p^2 + 2 \cdot n_m^2} \quad (4)$$

Note that our description of image formation mainly differs from experimental MP setups through non-scanned illumination, which is experimentally used to allow for widefield illumination without being limited by speckle-artifacts and to minimize the spatial extent of the molecular point-spread function.[3] Further details on the employed theoretical model, such as including the near-field effects of protein scattering at a refractive index interface, high NA focusing effects, and phase aberrations due to imaging into a layer of different refractive index (buffer vs glass) are given in Sections 4.1 and S1–S8.

We begin our numerical investigation with Hsp16.5, a highly symmetric small heat shock protein, which forms spherical 24-mers of m = 396 kDa adsorbed on a glass substrate covered by water to enable direct comparison with early experiments.[9] In a first iteration, we approximated the protein as a sphere of radius 5.6 nm and refractive index n_p = 1.480.[19] The substrate refractive index was chosen to be n_g = 1.515 as typical for the borosilicate microscope cover glass used in MP[18] but assumed to be atomically flat for simplicity. Experimentally, the limited full well capacity of CMOS imaging sensors requires both spatial and temporal summing of detected photoelectrons to reduce shot noise-induced background fluctuations and thereby optimize the attainable signal-to-noise ratio (SNR), taken as the ratio between the maximum signal amplitude introduced by the protein and the unknown background variations. In our simulations, we can take advantage of, in principle, unlimited full well capacities of the (virtual) camera pixels, which simplifies the image generation. The original experiments on Hsp16.5 used 2 × 2 spatial binning,

https://doi.org/10.1021/acsphotonics.3c00422
ACS Photonics XXXX, XXX, XXX–XXX

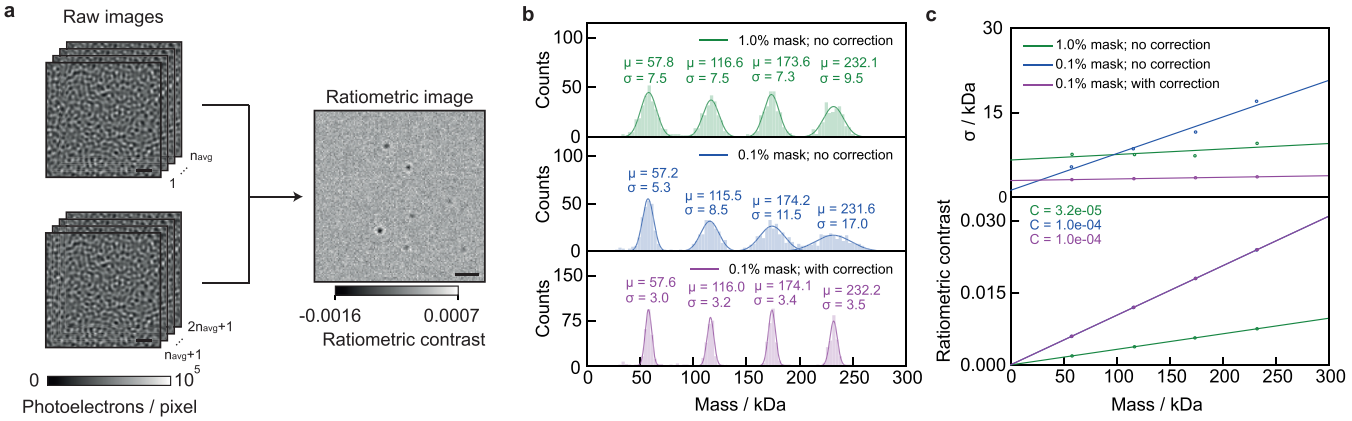


Figure 2. Simulation of protein landing events and resulting mass distributions as a function of mask strength. (a) Two consecutive sets of n_{avg} frames are averaged before computing the ratiometric image, revealing individual proteins landing at different positions as a function of time (scale bars: 1 μm). (b) Mass histograms for four oligomeric states of a protein simulated as spheres of radius 3.8 nm and refractive index 1.46 (representing BSA) using different mask strengths and local reflectivity correction.[9] In all cases, the power incident on the detector was kept constant. (c) Standard deviation of the fitted distributions (top) and ratiometric contrast (bottom) as a function of protein mass.

corresponding to 60 nm/pixel at the reported magnification,[9] followed by summing photoelectrons from 100 subsequent images, leading to a total of 10^7 detected photoelectrons per spatio-temporally binned camera pixel.

The quantity typically used to report the optical contrast in MP is that of the ratiometric contrast C

$$C = \frac{I_{det} - I_{bkg}}{I_{bkg}} = \frac{I_{det}}{I_{bkg}} - 1 \tag{5}$$

which is the relative difference between the measured intensity with (I_{det}) and without (I_{bkg}) the scatterer.[9] At the experimental illumination wavelength of λ = 445 nm, the simulated (here, without shot noise; details on simulation parameters are given in Section 4.2) ratiometric contrast of $C \sim 0.62\%$ is close to the experimental result of 0.6%, as is the appearance of Airy rings arising from plane wave illumination (Figure 1b). Note that the maximum contrast does not coincide with the nominal focus position (Figure 1c), instead requiring a displacement of the sample by ~190 nm along the optical axis to optimize the phase difference between scattered and reflected light by tuning the Gouy phase.[20] Small particles further away from the plane being imaged are typically not detectable in MP due to their weak scattering and rapid diffusion compared to the camera exposure time, yielding an even smaller signal per pixel as it is distributed over a larger spatial area once it has reached the image plane. Our simulated results suggest that our model produces image contrasts in good agreement with experiments, where the contrast is optimized by maximizing the (spatial) standard deviation of the glass roughness. At the same time, this agreement is to be taken with care, given the rather arbitrary definition of particle radius and refractive index.

To explore the dependence of the image contrast on the refractive index of both the medium and the substrate, which in principle are tuneable away from that of water (n_m = 1.333) and borosilicate glass (n_g = 1.515), we varied both parameters and evaluated the achievable SNR for a constant power density incident on the sample. While increasing both refractive indices for a fixed refractive index of the protein (n_p = 1.46), which we assume to be non-tuneable, leads to a modest increase in the achievable SNR (Figure 1d). Nevertheless, these considerations may be of interest for measurements in environments of different refractive indexes to water, such as those containing glycerol or sucrose, even though such refractive index tuning is unlikely to have a dramatic impact on the ultimate performance of light-based single-molecule detection and mass measurement.

MP requires the removal of a static background image, which is the main reason for generating ratiometric images.[1,2,21] This background resembles a speckle pattern generally attributed to nanoscale roughness of a microscope cover glass. Indeed, a recent report successfully correlated nanoscale roughness measured by atomic force microscopy with the corresponding image contrast.[22] Using the reported surface roughness parameters in terms of lateral (~100 nm) and vertical (~2 nm) dimensions, we simulated the resulting phase retardation ψ of the reflected field, through

$$I = |[\vec{E}_{ref} \cdot e^{i\cdot\psi} + \vec{E}_{sca}] \otimes \vec{h}|^2 \tag{6}$$

Note that ψ describes an effective phase change, which also accounts for the phase delay of the electric field that transmits through the glass−water interface and eventually leads to protein scattering. To obtain the speckle-like appearance in our simulation, we create a spatial array of randomly chosen numbers drawn from a uniform distribution $\mathcal{U}_{[-1,1]}$, which is then spatially low-pass filtered to yield a near diffraction-limited speckle pattern. Overall, this yields a surface height map Δh (varying between ±0.8 nm),[22] which can be converted into the relevant phase distortion (Figure 1e) via

$$\psi = \frac{2\pi}{\lambda} \cdot n_g \cdot \Delta h \tag{7}$$

Including this phase variation in our model predicts an image contrast in agreement with experimental results (Figure 1f),[9] effectively resulting in a locally varying reflection coefficient r.

Given that we can now produce raw images of both microscope cover glass and of single proteins with appropriate optical contrasts and spatial patterns, we can simulate a standard MP experiment (a general description of the typical experimental routine and materials is given in Sections 4.3 and 4.4). Here, a cleaned glass coverslip is usually covered by a dilute solution of biomolecules of interest, which bind non-specifically to the glass surface over time. These binding events

https://doi.org/10.1021/acsphotonics.3c00422
ACS Photonics XXXX, XXX, XXX−XXX

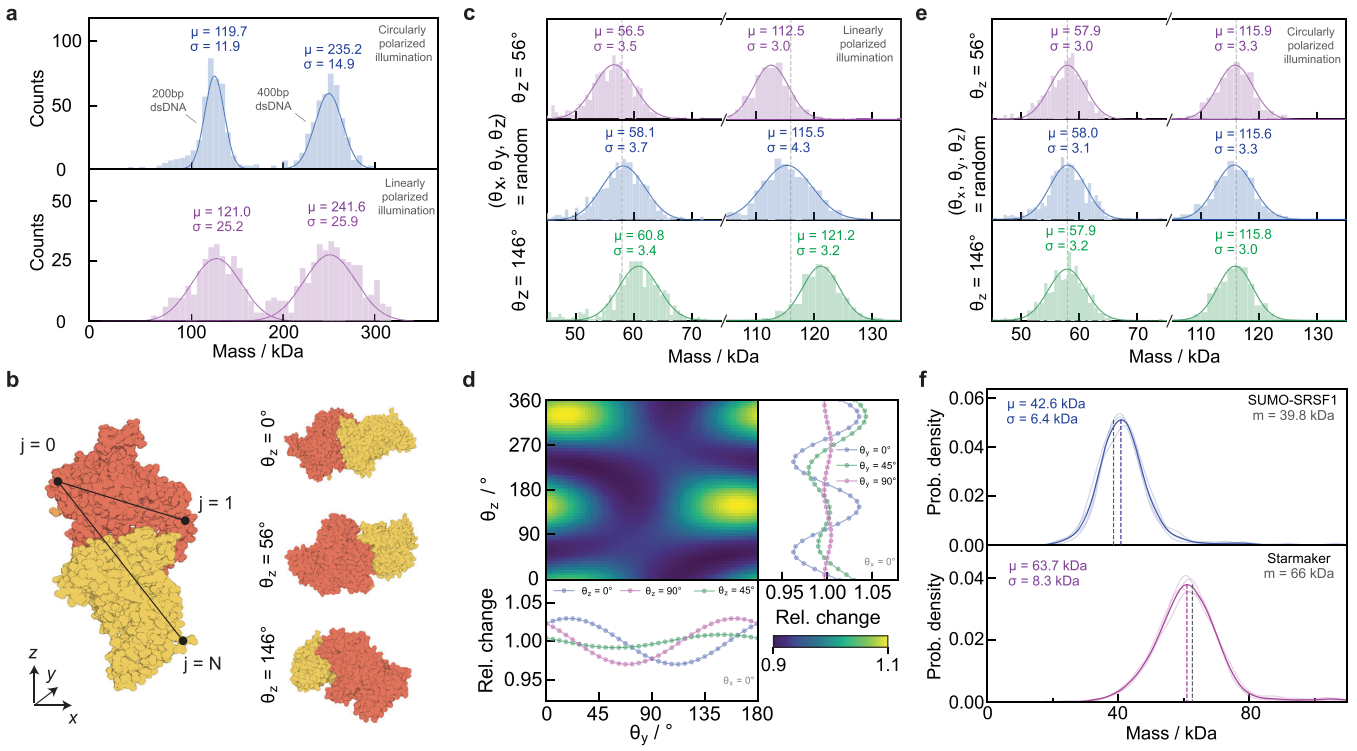


Figure 3. Dependence of mass measurement on biomolecular shape and illumination polarization. (a) Experimentally observed mass distributions for dsDNA illuminated with circularly (top) and linearly (bottom) polarized light. (b) Modeling the shape and orientation of a protein, here the dimer of BSA (PDB ID 3V03), by computing the corresponding polarizability tensor. (c) Mass histograms for a simulated landing assay where all BSA monomers (or dimers) land with the same (fixed) orientation ($\theta_z = 56°$; $\theta_z = 149°$), or with random orientations, while being illuminated with linearly polarized light. (d) Relative change of the ratiometric contrast for the BSA dimer for different orientations (changing θ_y and θ_z) relative to the incident polarization. (e) Same simulation as in (c), except for circularly polarized illumination. (f) Kernel density estimation (Gaussian; bandwidth = 2 kDa) of experimentally measured distributions of partially (SUMO-SRSF1; 4 repeats) and fully disordered (Starmaker; 4 repeats) proteins. The vertical dashed lines represent the mass inferred from the measurement (colored) and the expected mass (gray).

are best visualized by averaging a series of camera frames of the glass surface and computing the relative difference between consecutive sets as a function of time[3] (see Section 4.5). In this way, individual molecules are revealed, even though their contrast is much smaller than that generated by the glass roughness (Figure 2a).

Aside from detection sensitivity, the key performance parameter that determines the utility of MP is the achievable mass resolution, which originates in the measurement precision achievable on a molecule-by-molecule basis. Using illumination power (1% mask: 0.025 MW cm^{-2}; 0.1% mask: 0.25 MW cm^{-2}), wavelength (445 nm), and 10^7 detected photoelectrons per pixel (Section 4.2), subsequent to temporal and spatial binning, we find peak widths similar to optimal experimental results on the order of σ = 8−9 kDa (Figure 2b, top). Using a mask with lower transmission in combination with higher illumination power leads to reduced peak widths at low mass but a broadening as mass increases. This is caused by the influence of the glass roughness on the ratiometric contrast, which now not only depends on the scattering coefficient s but also on the locally varying reflectivity r of the glass−water interface (neglecting the purely scattering term)

$$C \sim 2 \cdot \frac{|s|}{|r|} \cdot \cos\varphi \tag{8}$$

with cos φ describing the phase difference between $\vec{E}_{\mathrm{ref}}$ and $\vec{E}_{\mathrm{sca}}$ at the detector plane. The glass roughness now results in the ratiometric contrast being dependent on where a particle lands on the glass coverslip. In the simulation, this broadening can be minimized by performing a correction step (based on ref 23) by multiplication with $\sqrt{I_{\mathrm{bkg}}}$, as this removes the dependence on r

$$C' = C \cdot \sqrt{I_{\mathrm{bkg}}} = 2 \cdot |s| \cdot \cos\varphi \cdot |E_{\mathrm{illu}}| \tag{9}$$

Note that applying eq 9 does retain most (~99%) of the overall landing events, while maintaining their respective landing coordinates (~91% matching in x, y, and frame number; see Section S17).

Applying this correction results in σ < 5 kDa peak widths but requires a homogeneous illumination field that is non-trivial to achieve in practice and the correction neglects an (unknown) phase contribution from the glass coverslip, which, in principle, results in residual broadening (see Section S6). In all cases, we observe behavior that agrees with expectations based on the selective reduction of reflected light by the transmission mask. A ten-fold reduction in reflected light is expected to increase the optical contrast by $10^{1/2}$, which would result in a 3.2-fold reduction in peak width (assuming a constant photon flux reaching the detector), comparable to our results (Figure 2c, top). Similarly, the concomitant increase in the contrast-to-mass conversion factor is also confirmed by our simulations (Figure 2c, bottom). These results validate the potentially high mass accuracy of MP, while assuming a perfectly spherical scatterer.

https://doi.org/10.1021/acsphotonics.3c00422
ACS Photonics XXXX, XXX, XXX−XXX

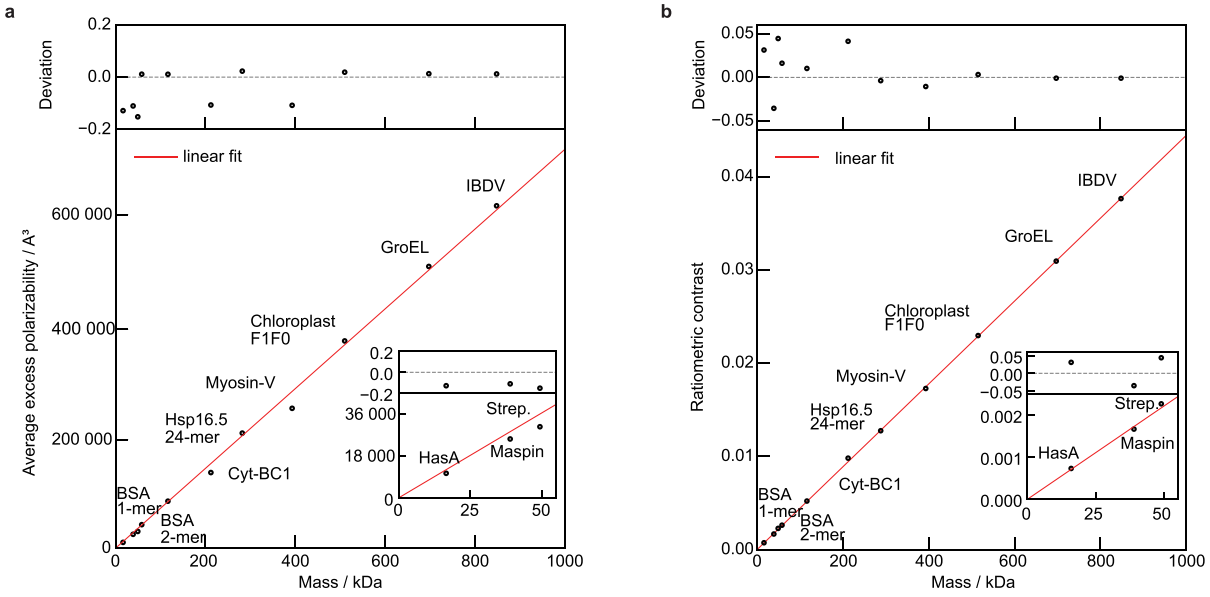


Figure 4. Mass scaling with molecular polarizability and image contrast. (a) Average excess polarizability and (b), calculated ratiometric contrast for proteins of mass 10−1000 kDa. Slope of linear fit (red line): 724 Å/kDa (a) and 4.4 × 10^{-5}/kDa (b). PDB IDs: HasA = 1B2V; Maspin = 1XQJ; Strep. = 4BX6; BSA = 3V03; Cyt-BC1 = 1BE3; Hsp16.5 = 1SHS; Myosin-V = 2DFS; Choroplast F1F0 = 6FKI; GroEL = 1GR5; IBDV = 2GSY.

To explore the effects of molecular shape and orientation beyond the spherical model, we chose double-stranded DNA (dsDNA), which effectively forms linear rods for a few hundred base pairs and below due to the persistence length of DNA.[24] We performed MP measurements of a mixture of dsDNA with different numbers of base pairs (Section 4.6) while illuminating the sample with circularly polarized light (Figure 3a, top) and found no significant effect of the elongated shape of the DNA molecule on the ratiometric contrast when compared to proteins of similar mass in terms of peak widths (simulated results, assuming a perfectly rod-like shape, indicate the same and are shown in Section S14). The slight broadening for 400 bp DNA likely stems from the fact that the length of the DNA (~136 nm) is no longer negligible compared to the diffraction limit, which can lead to the interferometric signal being spread over a slightly larger PSF and thus lower contrast, ultimately leading to peak broadening, an effect that becomes worse for longer DNA (see Sections 4.6 and S14). In principle, however, the enlarged PSF could be analyzed in a way to yield additional information on the size of the scatter, as shown in the supplement of Lee et al.,[25] where this was done to infer the orientation of gold nanorods. When performing the experiment with linearly polarized light (Figure 3a, bottom), we found a 2-fold increase in peak width, which stems from the underlying variability of contrasts measured on a molecule-by-molecule basis (also see Section S14). These results suggest that scatterer shape and orientation can play an important role when employing linearly polarized illumination coupled with fixed molecular orientations. At the same time, it shows that the use of circularly polarized light makes MP essentially insensitive to molecular shape.

While folded proteins do not exhibit the degree of anisotropy as short DNA strands, they are also not spherical, especially in the context of oligomerization. We therefore turned to a recently reported approach[26] using atomically resolved protein structures to compute the polarizability tensors of proteins from pairwise distances of all atoms in the molecule reported in the respective PDB structure (Figure 3b, left; also see Sections S7 and S8). The resulting polarizability tensor is a 3 × 3 matrix that encodes the anisotropic scattering of a particle as it connects any specific illumination direction with the scattering in all directions (denoted by x, y, and z; defined by the coordinate frame given in the PDB file)

$$\begin{pmatrix} E_{\text{scat}}{}^{x} \\ E_{\text{scat}}{}^{y} \\ E_{\text{scat}}{}^{z} \end{pmatrix} \propto \begin{bmatrix} \alpha_{1,1} & \alpha_{1,2} & \alpha_{1,3} \\ \alpha_{2,1} & \alpha_{2,2} & \alpha_{2,3} \\ \alpha_{3,1} & \alpha_{2,3} & \alpha_{3,3} \end{bmatrix} \cdot \begin{pmatrix} E_{\text{illu}}{}^{x} \\ E_{\text{illu}}{}^{y} \\ E_{\text{illu}}{}^{z} \end{pmatrix} \tag{10}$$

Note that any orientation of the protein can be achieved by multiplying the corresponding rotation matrices $\boldsymbol{R} = \boldsymbol{R}_z \cdot \boldsymbol{R}_y \cdot \boldsymbol{R}_x$ to determine the polarizability tensor $\boldsymbol{\alpha}_{\text{rot.}}$ in the new reference frame[27]

$$\boldsymbol{\alpha}_{\text{rot.}} = \boldsymbol{R}^T \cdot \begin{bmatrix} \alpha_{1,1} & \alpha_{1,2} & \alpha_{1,3} \\ \alpha_{2,1} & \alpha_{2,2} & \alpha_{2,3} \\ \alpha_{3,1} & \alpha_{2,3} & \alpha_{3,3} \end{bmatrix} \cdot \boldsymbol{R} \tag{11}$$

In the first instance, we simulated a series of landing events, where we fixed the orientation of BSA (major axis misaligned: θ_z = 56° or aligned: θ_z = 56 + 90° = 146°; with respect to the linearly polarized illumination), meaning that all landing events have the same, fixed, orientation and compared the resulting mass distributions to those obtained from randomly oriented molecules (with θ_x, θ_y, and θ_z chosen such that the orientation sampling is uniformly distributed in 3D), for both monomers and dimers (Figure 3b, right). We find significant deviations from the nominal mass in both cases, on the order of 4% of the expected mass for the monomer and dimer (Figure 3c). A deviation of 4% amounts to the maximum observed difference when sampling the full range of possible protein orientations (Figure 3d, with varying θ_y and θ_z). Repeating the simulation for BSA with circularly polarized light (Figure 3e) exhibited a drastically reduced dependence on protein orientation upon landing, now amounting to ≪1% of protein mass. While the

https://doi.org/10.1021/acsphotonics.3c00422
ACS Photonics XXXX, XXX, XXX−XXX

deviations observed for linearly polarized illumination lead to broadening of individual mass peaks, mass photometers reported to date largely rely on circularly polarized light,[3] making these measurements basically insensitive to protein shape. Note that the simulated protein mass for BSA of 58 kDa is lower than the mass based on its amino acid sequence (66 kDa) because the available PDB structure does not contain all atoms. When using the structure of BSA predicted by AlphaFold[28] (UniProt P02769), we find excellent agreement between the mass inferred from our simulation (64.3 kDa; shown in Section S13) and the AlphaFold mass (64.4 kDa).

In addition to protein orientation, the degree to which a protein is folded could also have a substantial effect on the relationship between optical contrast and mass through various factors such as amino acid density or the association of water and counterions, all of which affect the molecular polarizability. We therefore turned to partially and fully unfolded proteins and compared the measured mass using folded proteins as a mass calibrant to the expected mass (Section 4.7). The small ubiquitin-related modifier (SUMO)-tagged serine- and arginine-rich splicing factor 1[29] (SUMO-SRSF1) is a 39.8 kDa protein composed of an 11 kDA SUMO-tag, two 8−9 kDa structured domains (RRM1/2), and three intrinsically disordered domains totaling 10−11 kDa, making it 25% disordered by mass. The SUMO tag is a N-terminal carrier protein that promotes protein folding and stability, allowing for easier production of the desired protein.[30] Using a folded, oligomeric protein as a calibrant (dynamin-1 ΔPRD), we obtain a mass of 42.6 ± 0.64 kDa (Figure 3f, top; magenta vertical dashed line), in good agreement with the predicted mass (gray vertical dashed line) and within the error found for various folded proteins (Figure 4b). We then turned to Starmaker, a 66 kDa, fully disordered protein.[31] Due to its high negative overall charge, Starmaker does not bind adequately to standard microscope cover glass. We therefore functionalized the cover glass with (3-aminopropyl) triethoxysilane (APTES) to create a positively charged surface, obtaining a mass of 63.7 ± 0.4 kDa (Figure 3f, bottom; blue vertical dashed line), again in excellent agreement with the expected mass (gray vertical dashed line). In addition, in both cases we do not see any significant increase in the peak widths compared to folded proteins of similar mass.

We can now explore the previously reported linear relationship between optical contrast and mass for a variety of proteins, using random orientations for landing events and circularly polarized illumination as in the experiment. We find that the resulting (average) excess polarizability[16] scales linearly with the respective molecular mass, derived from the amino acid sequence in the respective PDB entry (Figure 4a; the respective PDB IDs are given in Section 4.8, Table 1), with the polarizability change per mass $\delta\alpha$

$$\alpha = \delta\alpha \cdot m = 724 \frac{\text{Å}^3}{\text{kDa}} \cdot m \tag{12}$$

which is slightly larger than that computed from bulk refractive index measurements[4] (460 Å^3 kDa^{-1}). This deviation might be (partially) explained by including the factor n_m^2 into the absolute polarizability value, whose definition depends on how the scattering process is introduced in the corresponding calculation. When computing α for a range of proteins, some exhibit substantial deviation (>15%) from the expected linear relationship, especially below 200 kDa. To determine whether these variations in polarizability are reflected in real MP measurements, we used the full polarizability tensor model to calculate the respective ratiometric contrast from simulated landing assay movies as a function of protein mass. We found a linear relationship, this time with an root-mean-square (rms) error of 2.4% and improved agreement above 200 kDa (Figure 4b), both of which agree well with experimental results.[3,9] The most likely reason for this improvement is that the average of the polarizability tensor does not report the scattering response to circularly polarized light but rather is an average measure of the particle scattering strength.[26] Note that by theoretically establishing the linear relationship between protein mass and optical contrast, we are now able to move away from the simplified description using relatively poorly defined parameters, such as the refractive index of a single-molecule n_p and its radius (implicitly assuming a spherical shape), toward a more proper characterization based on the structure of the protein. We emphasize that we used the expected mass as calculated from the atomic positions in the respective PDB file rather than the nominal protein mass for all calculations, meaning that most proteins are found below their mass expected from their amino acid sequence.

Equipped with a realistic and quantitative description of the biomolecular polarizability and image formation, we can now deduce a general equation of the achievable SNR as a function of the key experimental parameters. First, we derive the number of detected, scattered, photons (N_{sca}) per pixel and exposure time as (for the full derivation, see Section S10)

$$N_{sca} = \frac{8}{3} \cdot \pi^3 \cdot I_{illu} \cdot \frac{\Delta t \cdot \lambda}{h \cdot c} \cdot \frac{\text{OT} \cdot \text{QE} \cdot [1 - |r|^2] \cdot \mu \cdot \gamma}{\lambda^6} \cdot \frac{A_{BFP}}{f_{obj}^2} \cdot n_m^4 \cdot d_{px}^2 \cdot [\delta\alpha \cdot m]^2 \tag{13}$$

with the following parameters defined in SI base units: I_{illu}: illumination intensity [W/m^2]; Δt: effective exposure time per image [s]; λ: laser wavelength [m]; h: Planck's constant [J s]; c: speed of light [m/s] OT: optical throughput; QE: quantum efficiency; $|r|^2$: reflectivity of the glass−buffer interface; $\mu = \sin^{-1}(\min[\text{NA}/n_i, 1])/\pi$: collection efficiency of the detection objective; d_{px}: effective pixel size in sample space [m]; $A_{BFP} = \pi \cdot f_{obj}^2 \cdot (n_i \cdot n_m/n_g)^2$: area of the accessible BFP, limited by the critical angle [m^2]; γ: enhancement factor due to aplanatic factor and scattering beyond the critical angle (e.g., $\gamma \sim 1.58$, for a 1.42 NA oil-immersion objective; see Sections 4.9, Table 2 and S10); f_{obj}: focal length of detection objective [m]; n_m: refractive index of buffer medium; m: mass of protein [kDa]; $\delta\alpha$: polarizability change per kDa, i.e., the slope in Figure 4a.

The detected, reflected, number of photons N_{ref} (per pixel and exposure) is similarly given as

$$N_{ref} = I_{illu} \cdot \frac{\Delta t \cdot \lambda}{h \cdot c} \cdot \text{OT} \cdot \text{QE} \cdot |\tau|^2 \cdot |r|^2 \cdot d_{px}^2 \tag{14}$$

with $|\tau|^2$ indicating the (power) transmission coefficient of the mask in the BFP. The shot-noise limited SNR in terms of the ratiometric contrast follows then as (see Section S10)

$$\text{SNR} = \frac{2}{\sqrt{2}} \cdot \sqrt{N_{sca}} \tag{15}$$

with the factor of 2 originating from the interferometric nature of the signal and the $\sqrt{2}$ from the comparison of two subsequent sets of frames, required to remove the static background and form the ratiometric image. This can be

https://doi.org/10.1021/acsphotonics.3c00422
ACS Photonics XXXX, XXX, XXX−XXX

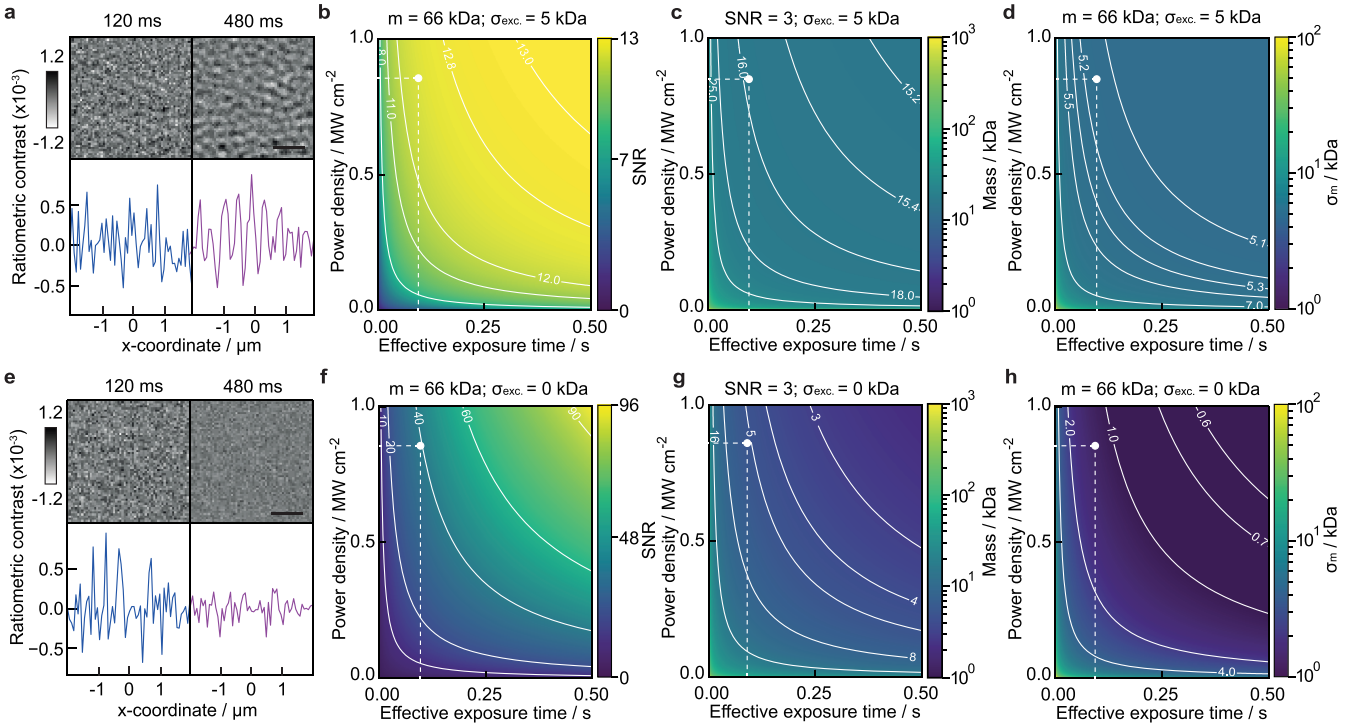


Figure 5. Current and future limits of optical mass measurement of single biomolecules. (a) Experimental ratiometric images of buffer medium only (scalebar = 1 μm) for different integration times. (b) Achievable SNR when detecting the BSA monomer (m = 66 kDa). (c) Smallest detectable mass $m_{q=3}$ at SNR = 3. (d) Mass resolution σ_m. All given as a function of effective exposure time and illuminating power density in the presence of excess noise on the order of 5 kDa. (e–h) Same (simulated) images and dependencies for purely shot noise limited performance. The white dots indicate experimental parameters from ref 3.

converted into an equivalent expression for the detection limit of MP, i.e., the smallest detectable mass m_q at a certain SNR level. For this, we set SNR = q and solve for m

$$\mathrm{SNR} = f(m) \quad \rightarrow \quad m_q = f(\mathrm{SNR} = q) \tag{16}$$

with f denoting a functional dependency. Similarly, we can define a measure describing the lowest achievable mass resolution σ_m, the Quantum-Cramer–Rao-Lower-Bound (QCRLB) as defined in ref 32

$$\sigma_m = \frac{m}{2}\cdot\frac{1}{\sqrt{N_{sca}}} = \frac{m}{\sqrt{2}}\cdot\frac{1}{\mathrm{SNR}} \tag{17}$$

Note that the QCRLB relates to the uncertainty introduced by the quantum nature of light itself, i.e., represents a fundamental limit that almost all optical measurement techniques will obey.

We also find that this optimum achievable mass resolution is directly linked to the SNR-equivalent mass, with $q = \sqrt{0.5}$ (see Section S10 for the derivation)

$$\sigma_m = m_{q=\sqrt{0.5}} \tag{18}$$

meaning that the highest attainable mass resolution is equivalent to the mass that achieves an SNR = $\sqrt{0.5}$ in the shot-noise limited regime.

Given that most of the parameters are essentially fixed in realistic experimental scenarios, or only vary marginally as a result of the details of experimental implementation, we can simplify these expressions to depend only on some key experimental details, specifically illumination power, exposure time, wavelength, and molecular mass (in kDa).

$$\mathrm{SNR} = K_1\cdot m\cdot\sqrt{\frac{I_{illu}\cdot\Delta t}{\lambda^5\cdot h\cdot c}} \tag{19}$$

with

$$K_1 = \sqrt{2}\cdot\sqrt{\frac{8}{3}\cdot\pi^3}\cdot\sqrt{\mathrm{OT}\cdot\mathrm{QE}}\cdot\sqrt{1-|r|^2}\cdot\sqrt{\mu}\cdot\sqrt{\gamma}$$
$$\cdot\frac{\sqrt{A_{BFP}}}{f_{obj}}\cdot n_m^2\cdot d_{px}\cdot\delta\alpha = 1.85\cdot10^{-33}$$

in SI units of m^4/kDa, assuming OT = 0.8; QE = 0.70; $|r|^2$ = 0.004; $|\tau|^2$ = 0.02; f_{obj} = 3 mm; $\delta\alpha$ = 724 Å^3/kDa; n_m = 1.333; μ = 0.40; γ = 1.58; A_{BFP} = 50 mm^2; d_{px} = 80 nm. Yielding an SNR ~ 21; $m_{q=3}$ ~ 9.5 kDa and σ_m ~ 2.2 kDa for I_{illu} = 0.1 MW cm^{-2}, Δt = 100 ms, λ = 445 nm, and m = 66 kDa.

Experimental images, however, including those consisting of buffer medium only, or even ultrapure water, reveal a dynamic, speckle-like background at high imaging sensitivity that cannot be removed by temporal averaging or attributed to sample drift, here shown by comparing 120 and 480 ms averaging time, plateauing at a contrast on the order of a 5 kDa protein (Figure 5a). This background is likely the current limiting factor to both improving mass resolution and the absolute detection limit of MP and is close to that reported recently using machine learning.[33] We currently have no clear explanation as to the origin of this additional noise source. As a first attempt we include it into our (SNR-) model, by adding it in quadrature to the shot-noise variance and find for the SNR including additional excess noise $\mathrm{SNR}_{exc.}$ (with $\sigma_{exc.}$ being the added variance in terms of photon counts)

https://doi.org/10.1021/acsphotonics.3c00422
ACS Photonics XXXX, XXX, XXX–XXX

$$\mathrm{SNR}_{\mathrm{exc.}} = \mathrm{SNR} \cdot \sqrt{\frac{2}{2 + \sigma_{\mathrm{exc.}}^2 \cdot N_{\mathrm{ref}}}} \xrightarrow{(I_{\mathrm{illu}}, \Delta t) \to (\infty, \infty)} \mathrm{SNR}_{\mathrm{exc.}}^{\mathrm{max}} = \frac{2}{\sigma_{\mathrm{exc.}}} \cdot \sqrt{K_2} \tag{20}$$

and the constant

$$K_2 = 8/3 \cdot \pi^3 \cdot |t_1|^2 \cdot \mu \cdot n_{\mathrm{m}}^4 \cdot \gamma \cdot A_{\mathrm{BFP}} \cdot \lambda^{-6} \cdot f_{\mathrm{obj}}^{-2} \cdot |\tau|^{-2} \cdot |r|^{-2} \cdot [\delta\alpha \cdot m]^2$$

Note that $\mathrm{SNR}_{\mathrm{exc.}}$ naturally results in shot-noise limited performance for $\sigma_{\mathrm{exc.}}^2 = 0$ and yields a finite, maximum $\mathrm{SNR}_{\mathrm{exc.}}^{\mathrm{max}}$. The added influence of $\sigma_{\mathrm{exc.}}$ has an overall impact on key performance parameters for protein detection and characterization, such as the achievable SNR (Figure 5b; with $\mathrm{SNR}_{\mathrm{exc.}}^{\mathrm{max}} \sim 13.2$), the lowest mass detectable (Figure 5c), and the achievable mass resolution (Figure 5d).

We can now compute theoretically achievable performance in terms of SNR, mass resolution, and detection limit (assumed for SNR = 3) both in the presence (Figure 5a–d) and absence (Figure 5e–h) of a non-shot noise contribution amounting to 5 kDa rms as currently observed experimentally. For realistic simulations, we find good agreement with previous reports,[3] such as an SNR of ∼12.5 for BSA at an exposure time of 100 ms and 0.85 MW cm^{-2}. Similarly, the recently reported SNR of 1.4 for a 9 kDa protein[33] appears realistic, given that we find $m_{q=1} \sim 5.2$ kDa with and ∼1.7 kDa without the additional baseline noise (for the same exposure time and illumination power density; see Section S16). Overall, our simulations in the absence of excess noise demonstrate that significant improvements in performance are still achievable with realistic illumination power (∼1 MW cm^{-2}) and exposure times (<s), such as 1 kDa mass resolution and few kDa detection sensitivity, even in the absence of advanced image processing.

3. DISCUSSION

We have presented a numerical approach that enables us to compute the optical contrast generated by individual biomolecules (smaller than the simulated pixel size) based on their atomic structure, orientation, and shape. Our results compare well with experimental data obtained by MP, suggesting that our model is indeed quantitative and asserting that there are no major (unknown) physical effects contributing to the current performance, which are not part of our theoretical description (such as Brownian motion or incomplete immobilization). We find a clear dependence of optical contrast on molecular shape in extreme cases such as short DNA strands illuminated by linearly polarized light, that weakens for folded proteins, and is essentially eliminated when using circular polarization. The predicted mass accuracy on the order of 2.4% rms matches that observed experimentally (2%), as does the computed dependence of optical contrast on the strength of the attenuation mask used. Our results on intrinsically disordered proteins support the hypothesis that the major determinants for the molecular polarizability are the constituent amino acids. In terms of achievable sensitivity, we present evidence for a dynamic, speckle-like background with a signal magnitude comparable to a 5 kDa protein. This background currently limits the ultimately achievable detection sensitivity and also affects the achievable mass resolution. The resolution is further affected by the static speckle-like background caused by microscope cover glass roughness, although it can be (partially) corrected for computationally.

Our results provide a quantitative framework for rationalizing label-free optical detection of single biomolecules. Polarizabilities, realistic incident power densities, and detection efficiencies effectively define achievable detection sensitivity and measurement precision at the single-molecule level, which translates into mass resolution. We emphasize that these relationships are independent of the optical approach, whether using total internal reflection[5] or scattering from nanochannels.[4] While there will be subtle differences in the achievable power densities and detection efficiencies, the presented limits are likely to be representative of what can be achieved in terms of mass measurement using light-based detection of single biomolecules. In all cases, when comparing experimental with these theoretical results, it is essential that any non-shot noise contributions to image background are considered and quantified carefully, given their influence on measurement sensitivity and precision.

What is most encouraging, however, is that there appears substantial scope for improvement that will enable quantitative characterization of complex mixtures of biomolecules with a resolution and sensitivity that covers almost all species and interactions. Moreover, implementation of approaches that enable extended observation of individual molecules either in nanochannels[4] or on bilayers[13,34] could bring about even further improvements, much in the spirit of the advances brought about by similar strategies in mass spectrometry, such as charge detection and orbitrap mass spectrometry.[35,36] Alternatively, if high resolution and sensitivity are not required, integration times can be drastically reduced, which will enable measurement at higher analyte concentrations, providing access to a broader range of affinities and ultimately weak interactions, further broadening the application scope of MP for characterizing biomolecular interactions and dynamics.

4. METHODS

4.1. Description of Image Formation. We model image formation in MP as a 3D (complex-valued) convolution of electric fields (reference and scattered) with an APSF,[27] which is implemented as a multiplication in Fourier space (through the convolution theorem). In contrast to previous reports on the computation of such an interferometric PSF,[15,16] we employ a Fourier-based approach (see "slice propagation method" in ref 37), with the benefits of being flexible and fast. This is mainly due to the fact that the necessary integrations over the BFP distribution (in refs 15 and 16), is replaced by fast-Fourier transforms (FFTs) while also enabling simple modifications of the complex amplitude transmission in the pupil of the imaging system (i.e., no radial symmetry required). Propagating the electric fields along the optical axis is achieved through the angular spectrum method,[38] where typical wrap-around effects are being avoided by replacing the FFT through a chirp-Z transform, as shown in ref 39. The strength of the reference and scattered fields are given by Fresnel's coefficients r and t (reflection and transmission) at a glass–water interface and by the scattering coefficient s of a small spherical particle, in the Rayleigh regime.[14] The scattering coefficient s, is related to the polarizability α, i.e., to the volume V of the spherical scatterer and the refractive indices of the protein n_{p} and the surrounding medium n_{m}. High NA focusing effects are included by taking into account the refraction of light toward the nominal focus,[27] the aplanatic factor,[37] and the additional Gouy phase shift of the scattered light[15] (all summarized in the $\vec{h}$-term). The influence of phase aberrations due to the

https://doi.org/10.1021/acsphotonics.3c00422
ACS Photonics XXXX, XXX, XXX−XXX

refractive index mismatch at the coverslip interface is described using the Gibson−Lanni model.[40] The scattering of the protein at the water−glass interface is given as a distribution at the back-focal plane of the objective,[41] which further enables us to add a mask that attenuates and/or delays the reference component.[9] A more detailed description of the underlying theoretical framework is given in Sections S1−S8.

4.2. Simulation Parameters. All numerical results assume a 1.42 NA oil-immersion (n_i = 1.515) objective while imaging at a wavelength of λ = 445 nm. The data shown in Figure 1 corresponds to an effective pixel size of 0.057 μm with 70 × 70 pixels, yielding a field-of-view (FoV) of 4 μm × 4 μm = 16 μm^2. The step size along the axial direction was chosen to match the lateral size, ±2 μm propagation from the nominal focus. In case of the simulated landing assay data (all remaining figures), the effective pixel size was 70 nm at 128 × 128 pixels, resulting in a FoV of ∼80 μm^2 while only computing a single in-focus slice. In this case, we set the attenuation mask to add a $\pi/2$ phase delay for $\vec{E}_{ref}$, which results in phase-matching of the two fields at the detector and optimum contrast at the nominal focal plane (see Sections S5 and S12). The illumination was set to be (right-) circularly polarized (simulated as the average of azimuthally distributed linear polarizations), except for Figure 3b,c where we defined it to be linearly polarized. In terms of illumination intensity $|\vec{E}_{illu}|^2$, we directly defined the detected photons at the detector, omitting the need to specify the overall efficiency of the detection system (including QE of the detector, losses at the optics, etc.). For the landing assay simulations, this yielded in 10^6 photons per 70 nm pixel per 500 μs exposure time, which corresponds to the 10^7 photoelectrons mentioned earlier, when performing the ratiometric calculation with N = 100. The simulations were run on a personal computer (Windows 10; Intel Core i7-6700 CPU @ 3.40 GHz; 16 GB RAM) and required ∼20 μs per simulated voxel (tested on a 256 × 256 × 256 grid).

4.3. General Measurement Routine. Data were either collected on a TwoMP (Refeyn, UK) or on a custom-built mass photometer. Coverslips (Menzel-Glaeser, 24 × 50 mm # 1.5 SPEZIAL; Thermo Fisher Scientific, U.S.) were cleaned to remove any contaminants by sonication for 5 min in 50/50 isopropanol and Milli-Q water (Merck, Germany), followed by 5 min in Milli-Q only. They were then dried using N_2 and stored in a covered box to prevent re-contamination. Immediately before measurement, a silicon gasket (Grace Bio-Labs CultureWell, 3 × 1 mm; U.S.) was laid on the coverslip to contain the sample. The coverslip was placed on a sample-stage (*xyz* for TwoMP; *z*-only for custom-built system) above the objective and a small amount of immersion oil (Zeiss, Immersol 518 F; Germany) was added between the coverslip and objective to form a continuous interface. Once the gasket and objective were aligned, buffer was loaded into the gasket using a micropipette (Gilson Pipetman, U.S.). In case of the TwoMP, this allowed the focus position of the setup to be found and the autofocus to be set before the protein began binding to the coverslip. The custom-built system was operated without an autofocus but proved to be stable enough over the recording time. Once focus had been set, the protein was diluted in an Eppendorf tube (Eppendorf, 1.5 mL; Germany) to give 20 μL of the sample, at a concentration of 10−50 nM. The sample was added to the gasket and the focus quickly rechecked. If the added sample had been kept on ice, the refractive index of the solution could change when the preloaded buffer and sample were mixed due to the difference in temperature, which changed the focus position slightly. A movie was then recorded. The protein dynamin-1 ΔPRD was used as a mass calibrant. It is highly stable, easily produced in large volumes, and oligomeric, providing a large number of species of known mass (sometimes up to 7) with which to calibrate, increasing the accuracy of the calibration.

4.4. Materials. Reagents used were from Sigma-Aldrich (U.S.), unless otherwise stated. Water was ultrapure Milli-Q (Merck, Germany), and all solutions were filtered through a 0.2 μm filter (Millipore, U.S.) before use.

The disordered protein SUMO-SRSF1 containing the solubilizing mutations Y37S and Y72S[42] was cloned by Gibson assembly into a pET28 plasmid (kind gift of B. B. Kragelund, University of Copenhagen, Denmark) downstream of a His_6-SUMO tag. The construct was transformed into chemically competent C41 *Escherichia coli* (Lucigen). Cultures were grown in 2x YT medium supplemented with 50 μg/mL kanamycin at 37 °C until an OD_{600} of ∼1.5 was reached, and protein expression was induced using 0.5 mM IPTG at 20 °C overnight. Cells were harvested, resuspended in buffer A (20 mM sodium phosphate pH 7.5, 800 mM NaCl, 5% glycerol, 0.01% Tween-20, 2 mM dithiothreitol (DTT), 150 mM L-arginine, and 150 mM L-glutamate) supplemented with 10 mM $MgCl_2$, 10 U/mL benzonase (Merck), and 20 μg/mL RNase A (NEB), and lysed by passing the suspension three times through a high-pressure homogenizer (HPL6, Maximator) cooled to 4 °C at 15,000−20,000 psi. The lysate was clarified by centrifugation and applied to a HisTrap Excel column (Cytiva, 5 mL per 1 L cell culture) equilibrated in buffer A. The column was washed with 15 column volumes (CVs) of buffer A, followed by 7 CVs buffer B (20 mM sodium phosphate pH 7.5, 3 M NaCl, 5% glycerol, 0.01% Tween-20, 2 mM DTT, 100 mM L-arginine, and 100 mM L-glutamate) and 7 CVs of 97% buffer A and 3% buffer C (20 mM sodium phosphate pH 7.5, 800 mM NaCl, 5% glycerol, 0.01% Tween-20, 2 mM DTT, 150 mM L-arginine, 150 mM L-glutamate, and 500 mM imidazole), before elution with buffer C. Fractions containing protein were pooled and diluted at least 8-fold with buffer D (20 mM HEPES pH 8.0, 10% glycerol, 0.001% Tween-20, 0.5 mM Tris-(2-carboxyethyl) phosphine, and TCEP) and 5 M NaCl until the sample was clear (∼0.8 M ionic strength or 55 mS/cm). Nucleotides and protein contaminants were removed from the sample by ion exchange chromatography using a MonoS 5/50 GL column (Cytiva) and a gradient of 30−70% of buffer E (20 mM HEPES pH 8.0, 2 M NaCl, 10% glycerol, 0.01% Tween-20, and 0.5 mM TCEP). Fractions with absorbance ratios of A260/A280 < 0.7 were pooled, flash frozen in liquid nitrogen, and stored at −80 °C. Starmaker was prepared as described previously.[31]

Stock solutions of SUMO-SRSF1 were at 29 μM protein. 1 μM aliquots were prepared in 20 mM HEPES (pH 7.4), 1 M NaCl, 1 mM DTT, and 5% glycerol buffer and flash frozen. High salt was required to stabilize the protein. DTT is a reducing agent, preventing oligomerization via disulfide bonds as SUMO-SRSF1 contains two internal cysteines. Starmaker was diluted in 20 mM Tris (pH 8.4) and 50 mM KCl. The concentration was unknown, so measurements of a range of dilutions from the original stock were taken to estimate the concentration. For Starmaker, coverslips were positively charged with APTES. Coverslips were cleaned as described before, then plasma cleaned for 8 min. The coverslips were

https://doi.org/10.1021/acsphotonics.3c00422
ACS Photonics XXXX, XXX, XXX−XXX

Table 1. PDB IDs for Different Proteins Used to Investigate the Linear Relationship between the Protein Mass and Scattering Strength (See Figure 4a,b)

Name	HasA	Maspin	Strep.	BSA	Cyt-BC1	Hsp16.5	Myosin-V	Chloro	GroEL	IBDV
PDB ID	1B2V	1XQJ	4BX6	3V03	1BE3	1SHS	2DFS	6FKI	1GR5	2GSY

washed in acetone and submerged in a 2% APTES/acetone solution. After 2 min, the coverslips were washed with acetone again and placed in an oven for 1.5 h at 110 °C. Finally, the coverslips were sonicated for 5 min in isopropanol, then water, and dried under N_2.

4.5. Data Analysis. For the analysis of the recorded data, we used an in-house python package. Raw frames from the measurement I_i were imported and converted into ratiometric frames, by averaging two stacks of n_{avg} raw frames; $\overline{I}_1$ from frame (i) to $(i + n_{avg})$ and $\overline{I}_2$ from frame $(i + n_{avg} + 1)$ to $(i + 2n_{avg} + 1)$

$$\overline{I}_1 = \frac{1}{n_{avg}} \cdot \sum_{l=i}^{i+n_{avg}} I_l, \qquad \overline{I}_2 = \frac{1}{n_{avg}} \cdot \sum_{l=i+n_{avg}+1}^{i+2n_{avg}+1} I_l \tag{21}$$

Those two stacks are then used to compute the relative difference

$$C = \frac{\overline{I}_2 - \overline{I}_1}{\overline{I}_2} \tag{22}$$

This eliminates the background signal (from the glass roughness) that is constant throughout the measurement. Protein binding events that occur during the measurement are not constant and hence are not removed when computing the relative difference. They appear as spots that fade in and out of the image as the protein binds and then becomes part of the background. Once the ratiometric frames have been calculated, protein binding events are identified by filtering the detected events for groups of pixels that meet a minimum radial symmetry and for pixels with a minimum signal above the background noise. A PSF (either a theoretical[3] or experimental model) is then fitted to each event to determine its contrast. The use of an experimental PSF is necessary when analyzing data that corresponds to illuminating the scatterer with linearly polarized light (e.g., in case of the DNA measurements). To obtain the experimental PSF, the initial PSF detection parameters (found with the theoretical model) are used to align the cropped images (typically 7 × 7 pixels) of the found landing events, by employing a cubic spline interpolation. Those sub-images are then averaged together, while outliers (Pearson's correlation test) are being removed from this average. The resulting cubic spline model is then used to determine the contrast of each landing event. The obtained contrast values for all events are then plotted as a mass histogram. For a particular species, the event contrasts are usually normally distributed allowing a Gaussian to be fitted to the respective peaks. The Gaussian position, width, and area were used to characterize the contrast of each peak μ, standard deviation σ, and counts, respectively. To generate a mass calibration, this procedure was applied to a measurement of dynamin-1 ΔPRD. The peak contrasts were used for calibration by plotting against the corresponding species mass. In case of the simulated data, the calibration was performed against the first four oligomeric states of BSA shown in Figure 2b while deliberately reducing the numerically applied shot noise to obtain an accurate calibration.

4.6. Measurements of Double-Stranded DNA. Data were taken on both a custom-built MP setup that uses 465 nm linearly polarized illumination and a TwoMP (Refeyn, UK) with 488 nm circularly polarized illumination. APTES coverslips with silicone gaskets were prepared via the procedure described above. 200, 400, and 600 base pair double-stranded DNA were prepared using standard procedures.[43] For the MP measurements 200, 400, and 600 bp dsDNA were diluted to 3, 4, and 4 nM, respectively, in phosphate buffered saline (PBS). 20 μL of the sample was added to a gasket containing 5 μL of buffer. Data were acquired for 120 s following sample refocusing. The contrast values were converted into mass using a calibration curve obtained from a measurement of dynamin-1 ΔPRD (adjusted by a factor ×1.25, as described in Section S14). Note that the 600 bp dsDNA was excluded from being presented in this work (see Figure 3a) due to its length being beyond the diffraction limit of the detection system, which yields additional mass broadening. For completeness, however, it is presented in Section S14.

For the custom-built linearly polarized MP setup, data were acquired with the following parameters: 959 μs exposure time, 787 fps, 3.4 × 11.7 μm^2 field of view, 3 × 3 pixel binning, and 4-fold temporal averaging. For the TwoMP, data were acquired with the following parameters: 1380 μs exposure time, 698 fps, 2.7 × 10.9 μm^2 field of view, 6 × 6 pixel binning, and 4-fold temporal averaging. Both datasets were analyzed such that the ratiometric window size amounted to ∼50 ms.

4.7. Measurements of Partially and Fully Unfolded Proteins. Data were taken using a TwoMP mass photometer (Refeyn, UK) and analyzed using a custom-written Python package, based on the procedure described in Young et al.[3] Coverslips (Menzel-Gläser, 24 × 50 mm # 1.5 SPEZIAL; Thermo Fisher Scientific, U.S.) were cleaned, a silicon gasket (Grace Bio-Labs CultureWell, 3 × 1 mm; U.S.) was laid on top, and 4 μL of buffer medium was added. Next, the protein was diluted in an Eppendorf tube (Eppendorf, 1.5 mL; Germany) to give 20 μL of sample and added into the gasket. Movies containing ∼1000−5000 binding events were recorded (60 s), analyzed, and converted into mass using a calibration curve (generated from a measurement of the oligomeric peaks of dynamin-1 ΔPRD).

SUMO-SRSF1 was diluted to 20 nM in a buffer of 20 mM HEPES (pH 7.4), 1 M NaCl, 1 mM DT, and 20 mM NaCl. A 20 mM Tris (pH 7.4) and 50 mM NaCl buffer was used to dilute Starmaker, allowing for a reduction or an increase in salt concentration upon measurement. For both proteins, 4 repeats were taken at each condition, with no significant unbinding in any repeat.

4.8. PDB IDs of Several Analyzed Proteins. A list containing PDB IDs of the respective proteins shown in Figure 4a,b is displayed in Table 1.

4.9. Enhancement Factor Describing the Scattering near a Glass−Water Interface. Table 2 reports the enhancement factor γ that describes the additionally detected scattering which corresponds to the near-field of a scatterer at

https://doi.org/10.1021/acsphotonics.3c00422
ACS Photonics XXXX, XXX, XXX−XXX

Table 2. Enhancement Factor Describing the Scattering of a Dipole at the Glass−Water Interface for Different NAs

NA	1.2	1.3	1.4	1.5
n_i	1.515	1.515	1.515	1.515
γ	1.14	1.21	1.53	1.67

the glass−water interface. Details on the computation of γ are given in Section S10.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsphotonics.3c00422.

Derivation of the underlying theoretical description and simulation routine (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Jan Becker − *The Kavli Institute for Nanoscience Discovery, University of Oxford, Oxford OX1 3QU, U.K.; Physical and Theoretical Chemistry Laboratory, Department of Chemistry, University of Oxford, Oxford OX1 3QZ, U.K.;* orcid.org/0000-0001-8529-4244; Email: jan.becker@chem.ox.ac.uk

Philipp Kukura − *The Kavli Institute for Nanoscience Discovery, University of Oxford, Oxford OX1 3QU, U.K.; Physical and Theoretical Chemistry Laboratory, Department of Chemistry, University of Oxford, Oxford OX1 3QZ, U.K.;* orcid.org/0000-0003-0136-7704; Email: philipp.kukura@chem.ox.ac.uk

Authors

Jack S. Peters − *The Kavli Institute for Nanoscience Discovery, University of Oxford, Oxford OX1 3QU, U.K.; Physical and Theoretical Chemistry Laboratory, Department of Chemistry, University of Oxford, Oxford OX1 3QZ, U.K.*

Ivor Crooks − *Physical and Theoretical Chemistry Laboratory, Department of Chemistry, University of Oxford, Oxford OX1 3QZ, U.K.*

Seham Helmi − *The Kavli Institute for Nanoscience Discovery, University of Oxford, Oxford OX1 3QU, U.K.; Physical and Theoretical Chemistry Laboratory, Department of Chemistry, University of Oxford, Oxford OX1 3QZ, U.K.*

Marie Synakewicz − *Department of Biochemistry, University of Zurich, Zurich 8057, Switzerland*

Benjamin Schuler − *Department of Biochemistry, University of Zurich, Zurich 8057, Switzerland; Department of Physics, University of Zurich, Zurich 8057, Switzerland;* orcid.org/0000-0002-5970-4251

Complete contact information is available at: https://pubs.acs.org/10.1021/acsphotonics.3c00422

Author Contributions

Conceptualization: J.B. and P.K.; methodology: J.B., J.S.P., I.C., and P.K.; investigation—simulation and theory: J.B.; DNA sample preparation: S.H.; intrinsically disordered proteins sample preparation: M.S.; investigation—experimental: J.S.P. and I.C.; data analysis: J.B.; visualization: J.B. and P.K.; funding acquisition: P.K., M.S., and B.S.; project administration: J.B. and P.K.; supervision: P.K. and B.S.; writing—original draft: J.B. and P.K.; writing—review and editing: J.B., J.S.P., I.C., S.H., M.S., B.S., and P.K.

Funding

This work was funded by the European Research Council (ERC) Consolidator Grant PHOTOMASS 819593, the Engineering and Physical Research Council (EPSRC) Leadership Fellowship EP/T03419X/1 (P.K.), a FEBS Long-term Fellowship (M.S.), and the Swiss National Science Foundation (B.S.).

Notes

The authors declare the following competing financial interest(s): P.K. is a nonexecutive director and shareholder of Refeyn Ltd., while I.C. is an employee of Refeyn Ltd. (his work has been carried out while being a student at University of Oxford). The other authors declare no competing interests. Data that supports the findings of this study are available from the authors upon request.

An initial version of this manuscript has been made available as the following pre-print: Jan Becker, Jack S. Peters, Ivor Crooks, Seham Helmi, Marie Synakewicz, Benjamin Schuler, Philipp Kukura; A quantitative description for optical mass measurement of single biomolecules; 2023; 2023.03.28.534430; bioRxiv; doi: https://doi.org/10.1101/2023.03.28.534430; https://www.biorxiv.org/content/10.1101/2023.03.28.534430v1.full (accessed 31st May 2023).

■ ACKNOWLEDGMENTS

We would like to thank Dan Loewenthal for creating the 3D rendering of the BSA dimer shown in Figure 3b and for generating the cover art, Roi Asor for fruitful discussions regarding obtaining the polarizability tensor from PDB structures, Magdalena Wojtas and Andrzej Ożyhar for the kind gift of the Starmaker sample, and Antoine Cléry and Andrea Holla for helpful discussions regarding protein preparation and selection, respectively.

■ REFERENCES

(1) Ortega Arroyo, J.; Andrecka, J.; Spillane, K. M.; Billington, N.; Takagi, Y.; Sellers, J. R.; Kukura, P. Label-Free, All-Optical Detection, Imaging, and Tracking of a Single Protein. *Nano Lett.* **2014**, *14*, 2065−2070.

(2) Piliarik, M.; Sandoghdar, V. Direct Optical Sensing of Single Unlabelled Proteins and Super-Resolution Imaging of Their Binding Sites. *Nat. Commun.* **2014**, *5*, 4495.

(3) Young, G.; Hundt, N.; Cole, D.; Fineberg, A.; Andrecka, J.; Tyler, A.; Olerinyova, A.; Ansari, A.; Marklund, E. G.; Collier, M. P.; Chandler, S. A.; Tkachenko, O.; Allen, J.; Crispin, M.; Billington, N.; Takagi, Y.; Sellers, J. R.; Eichmann, C.; Selenko, P.; Frey, L.; Riek, R.; Galpin, M. R.; Struwe, W. B.; Benesch, J. L. P.; Kukura, P. Quantitative Mass Imaging of Single Biological Macromolecules. *Science* **2018**, *360*, 423−427.

(4) Špačková, B.; Klein Moberg, H.; Fritzsche, J.; Tenghamn, J.; Sjösten, G.; Šípová-Jungová, H.; Albinsson, D.; Lubart, Q.; van Leeuwen, D.; Westerlund, F.; Midtvedt, D.; Esbjörner, E. K.; Käll, M.; Volpe, G.; Langhammer, C. Label-Free Nanofluidic Scattering Microscopy of Size and Mass of Single Diffusing Molecules and Nanoparticles. *Nat. Methods* **2022**, *19*, 751−758.

(5) Ma, G.; Wan, Z.; Yang, Y.; Zhang, P.; Wang, S.; Tao, N. Optical Imaging of Single-Protein Size, Charge, Mobility, and Binding. *Nat. Commun.* **2020**, *11*, 4768.

(6) Feyh, R.; Waeber, N. B.; Prinz, S.; Giammarinaro, P. I.; Bange, G.; Hochberg, G.; Hartmann, R. K.; Altegoer, F. Structure and Mechanistic Features of the Prokaryotic Minimal RNase P. *eLife* **2021**, *10*, No. e70160.

(7) Malay, A. D.; Miyazaki, N.; Biela, A.; Chakraborti, S.; Majsterkiewicz, K.; Stupka, I.; Kaplan, C. S.; Kowalczyk, A.; Piette, B. M. A. G.; Hochberg, G. K. A.; Wu, D.; Wrobel, T. P.; Fineberg, A.;

Kushwah, M. S.; Kelemen, M.; Vavpetič, P.; Pelicon, P.; Kukura, P.; Benesch, J. L. P.; Iwasaki, K.; Heddle, J. G. An Ultra-Stable Gold-Coordinated Protein Cage Displaying Reversible Assembly. *Nature* **2019**, *569*, 438−442.
(8) Bueno-Alejo, C. J.; Santana Vega, M.; Chaplin, A. K.; Farrow, C.; Axer, A.; Burley, G. A.; Dominguez, C.; Kara, H.; Paschalis, V.; Tubasum, S.; Eperon, I. C.; Clark, A. W.; Hudson, A. J. Surface Passivation with a Perfluoroalkane Brush Improves the Precision of Single-Molecule Measurements. *ACS Appl. Mater. Interfaces* **2022**, *14*, 49604−49616.
(9) Cole, D.; Young, G.; Weigel, A.; Sebesta, A.; Kukura, P. Label-Free Single-Molecule Imaging with Numerical-Aperture-Shaped Interferometric Scattering Microscopy. *ACS Photonics* **2017**, *4*, 211−216.
(10) Fineberg, A.; Surrey, T.; Kukura, P. Quantifying the Monomer−Dimer Equilibrium of Tubulin with Mass Photometry. *J. Mol. Biol.* **2020**, *432*, 6168−6172.
(11) Soltermann, F.; Struwe, W. B.; Kukura, P. Label-Free Methods for Optical in Vitro Characterization of Protein−Protein Interactions. *Phys. Chem. Chem. Phys.* **2021**, *23*, 16488−16500.
(12) Sonn-Segev, A.; Belacic, K.; Bodrug, T.; Young, G.; VanderLinden, R. T.; Schulman, B. A.; Schimpf, J.; Friedrich, T.; Dip, P. V.; Schwartz, T. U.; Bauer, B.; Peters, J.-M.; Struwe, W. B.; Benesch, J. L. P.; Brown, N. G.; Haselbach, D.; Kukura, P. Quantifying the Heterogeneity of Macromolecular Machines by Mass Photometry. *Nat. Commun.* **2020**, *11*, 1772.
(13) Foley, E. D. B.; Kushwah, M. S.; Young, G.; Kukura, P. Mass Photometry Enables Label-Free Tracking and Mass Measurement of Single Proteins on Lipid Bilayers. *Nat. Methods* **2021**, *18*, 1247−1252.
(14) Bohren, C. F.; Huffman, D. R. *Absorption and Scattering of Light by Small Particles*; John Wiley & Sons, 2008.
(15) Gholami Mahmoodabadi, R.; Taylor, R. W.; Kaller, M.; Spindler, S.; Mazaheri, M.; Kasaian, K.; Sandoghdar, V. Point Spread Function in Interferometric Scattering Microscopy (ISCAT) Part I: Aberrations in Defocusing and Axial Localization. *Opt. Express* **2020**, *28*, 25969.
(16) He, Y.; Lin, S.; Marc Louis Robert, H.; Li, H.; Zhang, P.; Piliarik, M.; Chen, X.-W. Multiscale Modeling and Analysis for High-Fidelity Interferometric Scattering Microscopy. *J. Phys. Appl. Phys.* **2021**, *54*, 274002.
(17) Goodman, J. W. *Introduction to Fourier Optics*; Roberts and Company Publishers, 2005.
(18) Saleh, B. E. A.; Teich, M. C. *Fundamentals of Photonics*; John Wiley & Sons, 2019.
(19) Guemouri, L.; Ogier, J.; Ramsden, J. J. Optical Properties of Protein Monolayers during Assembly. *J. Chem. Phys.* **1998**, *109*, 3265−3268.
(20) Kukura, P.; Celebrano, M.; Renn, A.; Sandoghdar, V. Imaging a Single Quantum Dot When It Is Dark. *Nano Lett.* **2009**, *9*, 926−929.
(21) Kukura, P.; Ewers, H.; Müller, C.; Renn, A.; Helenius, A.; Sandoghdar, V. High-Speed Nanoscopic Tracking of the Position and Orientation of a Single Virus. *Nat. Methods* **2009**, *6*, 923−927.
(22) Lin, S.; He, Y.; Feng, D.; Piliarik, M.; Chen, X.-W. Optical Fingerprint of Flat Substrate Surface and Marker-Free Lateral Displacement Detection with Angstrom-Level Precision. *Phys. Rev. Lett.* **2022**, *129*, 213201.
(23) Hantke, M.; Young, G. Interferometric Scattering Microscopy Methods and Systems. U.S. Patent 10,816,784 B1, October 27, 2020.
(24) Manning, G. S. The Persistence Length of DNA Is Reached from the Persistence Length of Its Null Isomer through an Internal Electrostatic Stretching Force. *Biophys. J.* **2006**, *91*, 3607−3616.
(25) Lee, I.-B.; Moon, H.-M.; Joo, J.-H.; Kim, K.-H.; Hong, S.-C.; Cho, M. Interferometric Scattering Microscopy with Polarization-Selective Dual Detection Scheme: Capturing the Orientational Information of Anisotropic Nanometric Objects. *ACS Photonics* **2018**, *5*, 797−804.
(26) Booth, L. S.; Browne, E. V.; Mauranyapin, N. P.; Madsen, L. S.; Barfoot, S.; Mark, A.; Bowen, W. P. Modelling of the Dynamic Polarizability of Macromolecules for Single-Molecule Optical Biosensing. *Sci. Rep.* **2022**, *12*, 1995.
(27) Novotny, L.; Hecht, B. *Principles of Nano-Optics*; Cambridge University Press, 2012 DOI: 10.1017/CBO9780511813535.
(28) Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; Bridgland, A.; Meyer, C.; Kohl, S. A. A.; Ballard, A. J.; Cowie, A.; Romera-Paredes, B.; Nikolov, S.; Jain, R.; Adler, J.; Back, T.; Petersen, S.; Reiman, D.; Clancy, E.; Zielinski, M.; Steinegger, M.; Pacholska, M.; Berghammer, T.; Bodenstein, S.; Silver, D.; Vinyals, O.; Senior, A. W.; Kavukcuoglu, K.; Kohli, P.; Hassabis, D. Highly Accurate Protein Structure Prediction with AlphaFold. *Nature* **2021**, *596*, 583−589.
(29) Das, S.; Krainer, A. R. Emerging Functions of SRSF1, Splicing Factor and Oncoprotein, in RNA Metabolism and Cancer. *Mol. Cancer Res.* **2014**, *12*, 1195−1204.
(30) Young, C. L.; Britton, Z. T.; Robinson, A. S. Recombinant Protein Expression and Purification: A Comprehensive Review of Affinity Tags and Microbial Applications. *Biotechnol. J.* **2012**, *7*, 620−634.
(31) Kapłon, T. M.; Rymarczyk, G.; Nocula-Ługowska, M.; Jakób, M.; Kochman, M.; Lisowski, M.; Szewczuk, Z.; Ożyhar, A. Starmaker Exhibits Properties of an Intrinsically Disordered Protein. *Biomacromolecules* **2008**, *9*, 2118−2125.
(32) Dong, J.; Maestre, D.; Conrad-Billroth, C.; Juffmann, T. Fundamental Bounds on the Precision of ISCAT, COBRI and Dark-Field Microscopy for 3D Localization and Mass Photometry. *J. Phys. Appl. Phys.* **2021**, *54*, 394002.
(33) Dahmardeh, M.; Mirzaalian Dastjerdi, H.; Mazal, H.; Köstler, H.; Sandoghdar, V. Self-Supervised Machine Learning Pushes the Sensitivity Limit in Label-Free Detection of Single Proteins below 10 KDa. *Nat. Methods* **2023**, *20*, 442−447.
(34) Heermann, T.; Steiert, F.; Ramm, B.; Hundt, N.; Schwille, P. Mass-Sensitive Particle Tracking to Elucidate the Membrane-Associated MinDE Reaction Cycle. *Nat. Methods* **2021**, *18*, 1239−1246.
(35) Wörner, T. P.; Snijder, J.; Bennett, A.; Agbandje-McKenna, M.; Makarov, A. A.; Heck, A. J. R. Resolving Heterogeneous Macromolecular Assemblies by Orbitrap-Based Single-Particle Charge Detection Mass Spectrometry. *Nat. Methods* **2020**, *17*, 395−398.
(36) Wörner, T. P.; Aizikov, K.; Snijder, J.; Fort, K. L.; Makarov, A. A.; Heck, A. J. R. Frequency Chasing of Individual Megadalton Ions in an Orbitrap Analyser Improves Precision of Analysis in Single-Molecule Mass Spectrometry. *Nat. Chem.* **2022**, *14*, 515−522.
(37) Miora, R. H. D.; Rohwer, E.; Kielhorn, M.; Sheppard, C. J. R.; Bosman, G.; Heintzmann, R. Calculating Point Spread Functions: Methods, Pitfalls and Solutions. arXiv January 31, **2023**. arXiv: 2301.13515.
(38) Matsushima, K.; Shimobaba, T. Band-Limited Angular Spectrum Method for Numerical Simulation of Free-Space Propagation in Far and Near Fields. *Opt. Express* **2009**, *17*, 19662−19673.
(39) Leutenegger, M.; Rao, R.; Leitgeb, R. A.; Lasser, T. Fast Focus Field Calculations. *Opt. Express* **2006**, *14*, 11277−11291.
(40) Gibson, S. F.; Lanni, F. Experimental Test of an Analytical Model of Aberration in an Oil-Immersion Objective Lens Used in Three-Dimensional Light Microscopy. *J. Opt. Soc. Am. A* **1992**, *9*, 154.
(41) Lieb, M. A.; Zavislan, J. M.; Novotny, L. Single-Molecule Orientations Determined by Direct Emission Pattern Imaging. *J. Opt. Soc. Am. B* **2004**, *21*, 1210−1215.
(42) Cléry, A.; Krepl, M.; Nguyen, C. K. X.; Moursy, A.; Jorjani, H.; Katsantoni, M.; Okoniewski, M.; Mittal, N.; Zavolan, M.; Sponer, J.; Allain, F. H.-T. Structure of SRSF1 RRM1 Bound to RNA Reveals an Unexpected Bimodal Mode of Interaction and Explains Its Involvement in SMN1 Exon7 Splicing. *Nat. Commun.* **2021**, *12*, 428.
(43) Wang, T.-Y.; Guo, L.; Zhang, J. Preparation of DNA Ladder Based on Multiplex PCR Technique. *J. Nucleic Acids* **2010**, *2010*, No. e421803.

https://doi.org/10.1021/acsphotonics.3c00422
ACS Photonics XXXX, XXX, XXX−XXX

Supplementary information:

A quantitative description for optical mass measurement of single biomolecules

Jan Becker[a,b,*], Jack S. Peters[a,b], Ivor Crooks[b], Seham Helmi[a,b],
Marie Synakewicz[c], Benjamin Schuler[c,d] and Philipp Kukura[a,b,*]

[a] The Kavli Institute for Nanoscience Discovery, University of Oxford, UK
[b] Physical and Theoretical Chemistry Laboratory, Departement of Chemistry, University of Oxford, UK
[c] Department of Biochemistry, University of Zurich, Switzerland
[d] Department of Physics, University of Zurich, Switzerland

[*] corresponding authors email: *jan.becker@chem.ox.ac.uk*
philipp.kukura@chem.ox.ac.uk

Contents

S1 General image formation model

We model the influence of the optical system (a *widefield* microscope) as a 3D convolution $\otimes$ with an amplitude response function $\mathbf{h}$. The detectable intensity I, is then given as [1]:

$$I(\mathbf{r}) = \left|\mathbf{E}_{ref} \otimes \mathbf{h}_{ref}(\mathbf{r}) + \mathbf{E}_{sca}(\mathbf{r}) \otimes \mathbf{h}_{sca}(\mathbf{r})\right|^2, \tag{S1}$$

with $\mathbf{E}_{ref}$ and $\mathbf{E}_{sca}$ being the reflected & scattered field at the nominal focal plane (= glass-water interface) and $\mathbf{r} = (x, y, z)^\top$ a vector representing spatial coordinates ($\mathbf{E}_{ref}$ is not spatially depending due to widefield illumination).
Note that the absolute square operation $|\mathbf{E}|^2$ is defined as:

$$|\mathbf{E}(\mathbf{r})|^2 = |E^{(1)}(\mathbf{r})|^2 + |E^{(2)}(\mathbf{r})|^2 + |E^{(3)}(\mathbf{r})|^2, \tag{S2}$$

with the superscript $^{(1),(2),(3)}$ denoting the electric field in the x, y and z direction. The convolution $\otimes$ is separately performed on each component of the vector fields, e.g. for the first component it is written as:

$$E^{(1)}(\mathbf{r}) \otimes h^{(1)}(\mathbf{r}) = \int_{-\infty}^{+\infty} d\mathbf{v}\ E^{(1)}(\mathbf{v}) \cdot h^{(1)}(\mathbf{r} - \mathbf{v}). \tag{S3}$$

Which is equivalent to a multiplication in *Fourier* space (*convolution* theorem [1]):

$$\mathcal{F}\left\{E^{(1)}(\mathbf{r}) \otimes h^{(1)}(\mathbf{r})\right\} = \tilde{E}^{(1)}(\mathbf{k}) \cdot \tilde{h}^{(1)}(\mathbf{k}), \tag{S4}$$

with $\mathbf{k} = (k_x, k_y, k_z)^\top$ denoting spatial frequencies and $\mathcal{F}$ being the Fourier transform, defined as:

$$\tilde{E}^{(1)}(\mathbf{k}) = \frac{1}{(2\pi)^{\dim\{d\mathbf{r}\}/2}} \int_{-\infty}^{+\infty} d\mathbf{r}\ E^{(1)}(\mathbf{r}) \cdot e^{i\mathbf{k}\cdot\mathbf{r}}, \tag{S5}$$

with $\dim\{d\mathbf{r}\}$ indicating the dimensionality of the Fourier transform (e.g. $d\mathbf{r} = (dx, dy)^\top \rightarrow \dim\{d\mathbf{r}\} = 2$).

Such a Fourier relationship connects the light distribution at the nominal focal plane (3D) with its counterpart in the back-focal-plane (*BFP*) of the microscope objective (2D only). When denoting the coordinates in the *BFP* with $\mathbf{r}' = (x', y')^\top$, we see that the focusing of a high *NA* objective, requires to account for the refraction of light towards the nominal focus. This is achieved by transforming cylindrical into spherical coordinates, which is given in [2] as:

$$\tilde{\mathbf{E}}(k_x, k_y, z=0) = \left[\tilde{\mathbf{E}}(\mathbf{k}') \cdot \begin{pmatrix}\cos\phi\\ \sin\phi\\ 0\end{pmatrix}\right] \cdot \begin{pmatrix}\cos\phi\cdot\cos\theta\\ \sin\phi\cdot\cos\theta\\ -\sin\theta\end{pmatrix} + \left[\tilde{\mathbf{E}}(\mathbf{k}') \cdot \begin{pmatrix}-\sin\phi\\ \cos\phi\\ 0\end{pmatrix}\right] \cdot \begin{pmatrix}-\sin\phi\\ \cos\phi\\ 0\end{pmatrix}, \tag{S6}$$

with ϕ and θ being the azimuthal and polar angle in the *BFP*, respectively. The first component of this sum is given as:

$$\left[\begin{pmatrix}\tilde{E}^{(1)}(\mathbf{k}')\\ \tilde{E}^{(2)}(\mathbf{k}')\\ 0\end{pmatrix} \cdot \begin{pmatrix}\cos\phi\\ \sin\phi\\ 0\end{pmatrix}\right] \cdot \begin{pmatrix}\cos\phi\cdot\cos\theta\\ \sin\phi\cdot\cos\theta\\ -\sin\theta\end{pmatrix} = \begin{pmatrix}\tilde{E}^{(1)}(\mathbf{k}')\cdot\cos^2\phi\cdot\cos\theta\ +\ \tilde{E}^{(2)}(\mathbf{k}')\cdot\sin\phi\cdot\cos\phi\cdot\cos\theta\\ \tilde{E}^{(1)}(\mathbf{k}')\cdot\sin\phi\cdot\cos\phi\cdot\cos\theta\ +\ \tilde{E}^{(2)}(\mathbf{k}')\cdot\sin^2\phi\cdot\cos\theta\\ -\tilde{E}^{(1)}(\mathbf{k}')\cdot\cos\phi\cdot\sin\theta\ -\ \tilde{E}^{(2)}(\mathbf{k}')\cdot\sin\phi\cdot\sin\theta\end{pmatrix}. \tag{S7}$$

And the second component as:

$$\left[\begin{pmatrix}\tilde{E}^{(1)}(\mathbf{k}')\\ \tilde{E}^{(2)}(\mathbf{k}')\\ 0\end{pmatrix} \cdot \begin{pmatrix}-\sin\phi\\ \cos\phi\\ 0\end{pmatrix}\right] \cdot \begin{pmatrix}-\sin\phi\\ \cos\phi\\ 0\end{pmatrix} = \begin{pmatrix}\tilde{E}^{(1)}(\mathbf{k}')\cdot\sin^2\phi\ -\ \tilde{E}^{(2)}(\mathbf{k}')\cdot\sin\phi\cdot\cos\phi\\ -\tilde{E}^{(1)}(\mathbf{k}')\cdot\sin\phi\cdot\cos\phi\ +\ \tilde{E}^{(2)}(\mathbf{k}')\cdot\cos^2\phi\\ 0\end{pmatrix}. \tag{S8}$$

Which can be summarized into a matrix equation, that relates a 2D *BFP* to a 3D (front-) focal plane field distribution:

$$\tilde{\mathbf{E}}(k_x, k_y, z=0) = \hat{\mathbf{T}}(\phi, \theta) \cdot \tilde{\mathbf{E}}(\mathbf{k}'), \tag{S9}$$

as we can write this as:

$$\begin{pmatrix}\tilde{E}^{(1)}(\mathbf{k}', z=0)\\ \tilde{E}^{(2)}(\mathbf{k}', z=0)\\ \tilde{E}^{(3)}(\mathbf{k}', z=0)\end{pmatrix} = \underbrace{\begin{bmatrix}\cos^2\phi\cdot\cos\theta + \sin^2\phi & \sin\phi\cdot\cos\phi\cdot(\cos\theta - 1) & 0\\ \sin\phi\cdot\cos\phi\cdot(\cos\theta - 1) & \sin^2\phi\cdot\cos\theta + \cos^2\phi & 0\\ -\cos\phi\cdot\sin\theta & -\sin\phi\cdot\sin\theta & 1\end{bmatrix}}_{\hat{\mathbf{T}}(\phi,\theta)} \cdot \begin{pmatrix}\tilde{E}^{(1)}(\mathbf{k}')\\ \tilde{E}^{(2)}(\mathbf{k}')\\ 0\end{pmatrix}. \tag{S10}$$

The real space representation $\mathbf{E}$, at $z = 0$, is then given as an inverse Fourier transform $\mathcal{F}^{-1}$ of each vector component (see *slice propagation method* in [3]):

$$\mathbf{E}(x, y, z = 0) = \mathcal{F}^{-1}\left\{\tilde{\mathbf{E}}(k_x, k_y, z = 0)\right\}, \tag{S11}$$

where the (inverse) Fourier transform is applied element wise and $\mathcal{F}^{-1}$ is the complex-conjugate operator of $\mathcal{F}$.

Next, we model the propagation along the optical axis z using the Gibson-Lanni model [4]. It describes the defocusing of an optical imaging system consisting of three layers (immersion, glass coverslip and sample, i.e. buffer). For this we define the optical path difference *OPD*, comparing the usage of the objective in the design case (immersion medium thickness $d_{i,*}$ and refractive index $n_{i,*}$) with the actual experimental scenario (denoted by d_i and n_i):

$$OPD(z, z_p) = \frac{n_m \cdot z_p}{\cos\theta_m} + \frac{n_i \cdot d_i(z, z_p)}{\cos\theta_i} - \left[\frac{n_{i,*} \cdot d_{i,*}}{\cos\theta_{i,*}} + n_i \cdot \sin\theta \cdot \left(z_p \cdot \tan\theta_m + d_i(z, z_p) \cdot \tan\theta_i - d_{i,*} \cdot \tan\theta_{i,*}\right)\right], \tag{S12}$$

with z_p the axial position of the protein, n_m the refractive index of the buffer medium (here water), n_g that of the glass coverslip and the polar angles θ in each layer which can be found using *Snell's* law, e.g. $\sin\theta_i \cdot n_i = \sin\theta_g \cdot n_g$. The z-dependence is introduced in terms of a variable immersion medium thickness d_i [4]:

$$d_i(z, z_p) = z - z_p + n_i \cdot \left(\frac{d_{i,*}}{n_{i,*}} - \frac{z_p}{n_m}\right). \tag{S13}$$

Finally, we obtain the full (3D), real space, electric field distribution by multiplying the *OPD* as a phase factor to the BFP-representation of the respective E-field, similar to the angular spectrum method [5]:

$$\mathbf{E}(\mathbf{r}) = \mathcal{F}^{-1}\left\{\mathcal{F}\{\mathbf{E}(x, y, z = 0)\} \;\cdot\; G(z, z_p)\right\}, \tag{S14}$$

with the phase factor G given as:

$$G(z, z_p) = e^{i \cdot \frac{2\pi}{\lambda} \cdot OPD(z, z_p)}. \tag{S15}$$

To avoid wrap-around artifacts, when propagating the fields (due to the use of the discrete Fourier transforms), we employ the *chirp* Z-transform *CZ* (or *zoomed* FFT) as described in [6], i.e. replacing the Fourier transforms in Eq. (S14) with *CZ*. In the following, however, we will still use the notation involving $\mathcal{F}$, as *CZ* is merely a tool to facilitate the artifact-free computation, when performing the numerical propagation of the electric fields.

S2 Strength of reflected & scattered fields

Both electric fields ($\mathbf{E}_{ref}$ and $\mathbf{E}_{sca}$) are related to the illumination field $\mathbf{E}_{illu}$ through the Fresnel reflection r, transmission t and the scattering coefficient s:

$$\tilde{\mathbf{E}}_{ref}(\mathbf{k}', z = 0) \;=\; r^{(\perp,\parallel)}(\theta) \;\cdot\; \tilde{\mathbf{E}}_{illu}(\mathbf{k}', z = 0), \tag{S16}$$

$$\tilde{\mathbf{E}}_{sca}(\mathbf{k}', z = 0) \;=\; s \cdot t_{1,2}^{(\perp,\parallel)}(\theta) \;\cdot\; \tilde{\mathbf{E}}_{illu}(\mathbf{k}', z = 0), \tag{S17}$$

with the subscript $_1$ and $_2$ indicating that the transmission through the glass-water interface has to be accounted for twice. Note that the Fresnel coefficients are defined with respect to $\perp$ & $\parallel$ polarization, i.e. we need to decompose $\mathbf{E}_{illu}$ accordingly:

$$\begin{pmatrix} \tilde{E}_{illu}^{(\perp)}(\mathbf{k}', z = 0) \\ \tilde{E}_{illu}^{(\parallel)}(\mathbf{k}', z = 0) \\ 0 \end{pmatrix} = \underbrace{\begin{bmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{\hat{\mathbf{P}}(\phi)} \cdot \begin{pmatrix} \tilde{E}_{illu}^{(1)}(\mathbf{k}', z = 0) \\ \tilde{E}_{illu}^{(2)}(\mathbf{k}', z = 0) \\ 0 \end{pmatrix}. \tag{S18}$$

Here we have assumed that $\mathbf{E}_{illu}$ is a perfect transverse wave, i.e. fully unpolarized along z. In our simulation, $\mathbf{E}_{illu}$ is given as a single plane wave (due to widefield illumination), traveling at $\theta = 0^\circ$ with a certain polarization, e.g. here x'-polarized.

The scattering coefficient s is related to the scattering cross section σ_{sca} through:

$$s = \sqrt{\mu \cdot \frac{\sigma_{sca}}{A}}, \tag{S19}$$

with $\mu = \arcsin\{\min(NA/n_i, 1)\}$ being the collection efficiency of the objective (*NA* - numerical aperture; n_i refractive index of immersion) [7] and A the area that samples the photon flux, i.e. the effective pixel size of the detector.
The scattering cross section σ_{sca}, assuming that the protein size is much smaller than λ, is given as [8]:

$$\sigma_{sca} = \frac{1}{6\pi} \cdot \left(\frac{2\pi}{\lambda} \cdot n_m \right)^4 \cdot |\alpha|^2, \tag{S20}$$

with λ being the vacuum wavelength of light and α the *polarizability* of the protein. In a first approximation we treat the protein as a spherical particle of homogeneous refractive index n_p, for which an analytical expression for α can be found [8]:

$$\alpha = 3 \cdot V_{sphere} \cdot \frac{n_p^2 - n_m^2}{n_p^2 + 2 \cdot n_m^2}, \tag{S21}$$

with V_{sphere} being the volume of the respective spherical particle with radius a:

$$V_{sphere} = \frac{4}{3}\pi \cdot a^3. \tag{S22}$$

S3 Amplitude point-spread functions including near-field components

Next, we compute the amplitude response $\mathbf{h}_{ref}$ by defining the optical systems band limit in Fourier space:

$$\tilde{h}^{(1)}(\mathbf{k}', z = 0) = \begin{cases} 1 & |\mathbf{k}'| \leq k_{max} \\ 0 & \text{else} \end{cases}, \tag{S23}$$

with k_{max} being the maximum transferable spatial frequency given by Abbe's diffraction limit [9] in the fully *coherent* case:

$$k_{max} = \frac{2\pi}{\lambda} \cdot NA. \tag{S24}$$

By making use of Eq. (S9), (S11), (S14) & $z_p = 0$ (nominal focal plane), we find the corresponding real space representation:

$$\mathbf{h}_{ref}(\mathbf{r}) = \frac{-i}{\sqrt{\cos\theta}} \cdot \mathcal{F}^{-1}\left\{ \hat{\mathbf{T}}(\phi, \theta) \cdot G(z, 0) \cdot \tilde{\mathbf{h}}_{ref}(\mathbf{k}', z = 0) \right\}. \tag{S25}$$

With $\sqrt{\cos\theta}$ describing the *aplanatic* factor [2] and $-i$ a phase shift associated with the Huygens-Fresnel principle [10].

Similarly we can compute the amplitude response of the scattering $\mathbf{h}_{sca}$, which can be modeled as the presence of an induced dipole in the protein. Hence, the detectable light is governed by the interaction of such a dipole at a distance δ from the refractive index interface. Here we adopt the work of Lieb et al. [11], who describe the electric field distribution of such a dipole at an interface in the *BFP* of the microscope objective. The corresponding fields, in $\perp$ & $\parallel$ polarization, are given as:

$$\begin{aligned} \tilde{h}_{sca}^{(\parallel)}(\mathbf{k}', z = 0) &= c_1(\theta) \cdot \sin\theta \cdot \cos\Theta \\ &+ c_2(\theta) \cdot \cos\theta \cdot \sin\Theta \cdot \cos(\phi - \Phi), \end{aligned} \tag{S26}$$

$$\tilde{h}_{sca}^{(\perp)}(\mathbf{k}', z = 0) = c_3(\theta) \cdot \sin\Theta \cdot \sin(\phi - \Phi), \tag{S27}$$

with Θ and Φ the azimuth and polar angle of the dipole axis. The constants $c_1 - c_3$ are given according to:

$$c_1(\theta) = n^2 \cdot \frac{\cos\theta}{\cos\theta_s} \cdot t_2^{(\parallel)}(\theta_s) \cdot \Pi(\theta_s), \tag{S28}$$

$$c_2(\theta) = n \cdot t_2^{(\parallel)}(\theta_s) \cdot \Pi(\theta_s), \tag{S29}$$

$$c_3(\theta) = -n \cdot \frac{\cos\theta}{\cos\theta_s} \cdot t_2^{(\perp)}(\theta_s) \cdot \Pi(\theta_s), \tag{S30}$$

$$\Pi(\theta_s) = \exp(i\tfrac{2\pi}{\lambda} n_g \cos\theta_s \delta), \tag{S31}$$

where $n = n_m/n_g$ and $\theta_s = \arcsin(n \cdot \sin\theta)$.
Finally, we compute the amplitude response $\mathbf{h}_{sca}$ by setting $z_p = a$ as:

$$\mathbf{h}_{sca}(\mathbf{r}) = \frac{1}{\sqrt{\cos\theta}} \cdot \mathcal{F}^{-1}\left\{ \hat{\mathbf{T}}(\phi, \theta) \cdot \hat{\mathbf{P}}(\phi) \cdot G(z, a) \cdot \tilde{\mathbf{h}}_{sca}(\mathbf{k}', z = 0) \right\}, \tag{S32}$$

where we already account for the *Gouy* phase shift [12].

S4 Information content in the interferometric & ratiometric images

Finally we can express Eq. (S1) in terms of the *imaged* electric fields $\mathbf{A}$, i.e. as they appear on the detector:

$$I(\mathbf{r}) = |\mathbf{A}_{ref}|^2 + |\mathbf{A}_{sca}(\mathbf{r})|^2 + 2 \cdot |\mathbf{A}_{ref}| \cdot |\mathbf{A}_{sca}(\mathbf{r})| \cdot \cos\varphi(\mathbf{r}), \tag{S33}$$

with φ being the phase difference between the scattered and reflected fields at the detector plane, which are given as:

$$\mathbf{A}_{ref} = \mathbf{E}_{ref} \otimes \mathbf{h}_{ref}(\mathbf{r}), \tag{S34}$$

$$\mathbf{A}_{sca}(\mathbf{r}) = \mathbf{E}_{sca}(\mathbf{r}) \otimes \mathbf{h}_{sca}(\mathbf{r}). \tag{S35}$$

To access the interferometric part of the signal we compute the *ratiometric* contrast C, according to [13]:

$$C(\mathbf{r}) = \frac{I(\mathbf{r}) - I_{bkg}}{I_{bkg}}, \tag{S36}$$

with I_{bkg} being the detectable intensity without the protein:

$$I_{bkg} = |\mathbf{A}_{ref}|^2. \tag{S37}$$

Assuming that the scattered field is much weaker than the reflected counterpart, i.e. $|\mathbf{A}_{ref}| \gg |\mathbf{A}_{sca}(\mathbf{r})|$, we get:

$$C(\mathbf{r}) \approx |\mathbf{A}_{sca}(\mathbf{r})| \cdot \frac{2 \cdot \cos\varphi(\mathbf{r})}{|\mathbf{A}_{ref}|}. \tag{S38}$$

Showing that the detectable contrast is proportional to the scattered field on the detector, which itself is proportional to the polarizability of the protein. The sensitivity of the detection can be increased either by optimizing the phase difference φ (such that $\cos\varphi = 1$, i.e. phase matching conditions) or by attenuating the reference field (using a mask in the *BFP*, see [14]).

S5 Increasing contrast by attenuating & phase shifting the reference field

In principle mass photometry is only limited by shot noise, i.e. given by the number of detectable scattered photons. In practice we are limited by the full-well-depth of the detector, which for small proteins is mostly depleted by the reference field. However, as Eq. (S38) suggests, we can increase the sensitivity of the detection by decreasing $|\mathbf{A}_{ref}|$. Experimentally this is achieved by placing a mask in the *BFP* of the microscope objective [14]. This is where the scattered and reference fields are separable in terms of the spatial frequency components. We model this by multiplying $\tilde{\mathbf{h}}$ (both: the reference & scattered field) with a Fourier mask $\tilde{\tau}$, given as:

$$\tilde{\tau}(\mathbf{k}') = \begin{cases} |\tilde{\tau}| \cdot e^{i \cdot \arg\{\tilde{\tau}\}} & |\mathbf{k}'| \leq k_{mask} \\ 0 & \text{else} \end{cases}, \tag{S39}$$

with $|\tilde{\tau}|$ being the transmissivity and $\arg\{\tilde{\tau}\}$ the phase shift of the mask. k_{mask} is related to the spatial extent of the mask with respect of the *BFP* diameter:

$$k_{mask} = \frac{R_{mask}}{f_{obj} \cdot NA}, \tag{S40}$$

with R_{mask} the radius of the mask in spatial units (e.g. mm).
The results shown in this publication correspond to a mask size that corresponds to an effective *NA* of 0.58. The term *phase matched* refers to an additional phase shift that the mask introduces which was set to $\arg\{\tilde{\tau}\} = \pi/2$, yielding an optimum contrast at the nominal focal plane $z = 0$ (see Fig. S1 b bottom).

S6 Glass roughness due to phase variations in the reference field

When dealing with real experimental data one typically observes a large, speckle-like, background in the raw data. Following the findings of [15], we modify the constant reference field at the nominal focal plane by a phase variation Ψ.

$$I(\mathbf{r}) = \left| \left[\mathbf{E}_{ref} \cdot e^{i\Psi(\mathbf{r})} \right] \otimes \mathbf{h}_{ref}(\mathbf{r}) + \mathbf{E}_{sca}(\mathbf{r}) \otimes \mathbf{h}_{sca}(\mathbf{r}) \right|^2. \tag{S41}$$

This phase distortion is due to the roughness of the glass coverslip. Note that Ψ describes an *effective* phase change, which also accounts for the phase delay of the electric field that transmits through the glass-water interface and eventually leads to the protein scattering. In case of small phase variations, we can approximate this exponential as:

$$e^{i\Psi(\mathbf{r})} \approx 1 + i\Psi(\mathbf{r}). \tag{S42}$$

With which we rewrite Eq. (S41) into:

$$I(\mathbf{r}) \approx \left|\mathbf{A}_{ref} + \mathbf{A}_{glass}(\mathbf{r}) + \mathbf{A}_{sca}(\mathbf{r})\right|^2, \tag{S43}$$

with $\mathbf{A}_{glass}$ being the complex field (at the detector) associated to the glass roughness.

$$\mathbf{A}_{glass}(\mathbf{r}) = \left[i \cdot \mathbf{E}_{ref} \cdot \Psi(\mathbf{r})\right] \otimes \mathbf{h}_{ref}(\mathbf{r}), \tag{S44}$$

with ψ being the phase difference between the fields due to glass roughness and protein scattering.
The detected intensity is then given as:

$$\begin{aligned} I(\mathbf{r}) &= |\mathbf{A}_{ref} + \mathbf{A}_{glass}(\mathbf{r}) + \mathbf{A}_{sca}(\mathbf{r})|^2 = \\ &= |\mathbf{A}_{ref}|^2 + |\mathbf{A}_{glass}(\mathbf{r})|^2 + |\mathbf{A}_{sca}(\mathbf{r})|^2 \\ &+ 2 \cdot |\mathbf{A}_{ref}| \cdot |\mathbf{A}_{glass}(\mathbf{r})| \cdot \cos\psi(\mathbf{r}) \\ &+ 2 \cdot |\mathbf{A}_{ref}| \cdot |\mathbf{A}_{sca}(\mathbf{r})| \cdot \cos\varphi(\mathbf{r}) \\ &+ 2 \cdot |\mathbf{A}_{glass}(\mathbf{r})| \cdot |\mathbf{A}_{sca}(\mathbf{r})| \cdot \cos[\psi(\mathbf{r}) - \varphi(\mathbf{r})]. \end{aligned} \tag{S45}$$

The ratiometric signal Eq. (S36) is now given as:

$$C(\mathbf{r}) \approx |\mathbf{A}_{sca}(\mathbf{r})| \cdot 2 \cdot \left[\frac{|\mathbf{A}_{ref}| \cdot \cos\varphi(\mathbf{r})}{|\mathbf{A}_{ref} + \mathbf{A}_{glass}(\mathbf{r})|^2} + \frac{|\mathbf{A}_{glass}(\mathbf{r})| \cdot \cos[\psi(\mathbf{r}) - \varphi(\mathbf{r})]}{|\mathbf{A}_{ref} + \mathbf{A}_{glass}(\mathbf{r})|^2}\right]. \tag{S46}$$

Alternatively, we could also write the intensity I in terms of an auxiliary reference field $\mathbf{A}'_{ref} = \mathbf{A}_{ref} + \mathbf{A}_{glass}$:

$$I(\mathbf{r}) = |\mathbf{A}'_{ref}(\mathbf{r}) + \mathbf{A}_{sca}(\mathbf{r})|^2. \tag{S47}$$

Which we now express with respect to an overall phase shift ζ, comparing the field scattered from the protein with that of the total reference field (including the phase variation due to the glass).

$$\begin{aligned} I(\mathbf{r}) &= |\mathbf{A}'_{ref}(\mathbf{r})|^2 + |\mathbf{A}_{sca}(\mathbf{r})|^2 + 2 \cdot |\mathbf{A}'_{ref}(\mathbf{r})| \cdot |\mathbf{A}_{sca}(\mathbf{r})| \cdot \cos\zeta(\mathbf{r}) = \\ &= |\mathbf{A}_{ref} + \mathbf{A}_{glass}(\mathbf{r})|^2 + |\mathbf{A}_{sca}(\mathbf{r})|^2 + 2 \cdot |\mathbf{A}_{ref} + \mathbf{A}_{glass}(\mathbf{r})| \cdot |\mathbf{A}_{sca}(\mathbf{r})| \cdot \cos\zeta(\mathbf{r}) = \\ &= |\mathbf{A}_{ref}|^2 + |\mathbf{A}_{glass}(\mathbf{r})|^2 + |\mathbf{A}_{sca}(\mathbf{r})|^2 + 2 \cdot |\mathbf{A}_{ref}| \cdot |\mathbf{A}_{glass}(\mathbf{r})| \cdot \cos\psi(\mathbf{r}) \\ &+ 2 \cdot |\mathbf{A}_{ref} + \mathbf{A}_{glass}(\mathbf{r})| \cdot |\mathbf{A}_{sca}(\mathbf{r})| \cdot \cos\zeta(\mathbf{r}). \end{aligned} \tag{S48}$$

Which helps us to identify that:

$$\begin{aligned} |\mathbf{A}_{ref}| \cdot \cos\varphi(\mathbf{r}) + |\mathbf{A}_{glass}(\mathbf{r})| \cdot \cos[\psi(\mathbf{r}) - \varphi(\mathbf{r})] \\ = |\mathbf{A}_{ref} + \mathbf{A}_{glass}(\mathbf{r})| \cdot \cos\zeta(\mathbf{r}). \end{aligned} \tag{S49}$$

Yielding the following for the ratiometric contrast:

$$C(\mathbf{r}) \approx |\mathbf{A}_{sca}(\mathbf{r})| \cdot \frac{2 \cdot \cos\zeta(\mathbf{r})}{|\mathbf{A}_{ref} + \mathbf{A}_{glass}(\mathbf{r})|}. \tag{S50}$$

Note that this makes the ratiometric contrast vary with the landing position of the particles, which itself leads to mass broadening. A (partial) correction of this can be achieved by computing the following (based on [16]):

$$C'(\mathbf{r}) = C(\mathbf{r}) \cdot \sqrt{I_{bkg}(\mathbf{r})}, \tag{S51}$$

with the background field given as:

$$I_{bkg}(\mathbf{r}) = \left|\mathbf{A}_{ref} + \mathbf{A}_{glass}(\mathbf{r})\right|^2. \tag{S52}$$

This yields the *corrected* ratiometric signal:

$$C'(\mathbf{r}) \approx |\mathbf{A}_{sca}(\mathbf{r})| \quad \cdot \quad 2 \cdot \cos\zeta(\mathbf{r}). \tag{S53}$$

This quantity (in principle) still depends on the landing position of the particle, as $\cos\zeta$ is spatially varying due to the glass roughness. Nevertheless, this influence is much weaker (as $|\cos\zeta| \leq 1$). Another practical limitation is the fact that $|\mathbf{A}_{sca}| \propto |\mathbf{E}_{illu}|$, i.e. any inhomogeneous illumination profile will also lead in mass broadening. Hence, proper flat fielding is required, as already discussed in [16].

S7 Scattering anisotropy of an ellipsoidal particle

When treating *anisotropy* in the polarizability of a protein, we need to reformulate Eq. (S20) by replacing the polarizability scalar α with the polarizability tensor $\hat{\alpha}$. In case of shape-dependent anisotropy (e.g. elliptical shape of the protein), the polarizability tensor $\hat{\alpha}$ is defined according to [8]:

$$\hat{\alpha} = \begin{bmatrix} \alpha_a & 0 & 0 \\ 0 & \alpha_b & 0 \\ 0 & 0 & \alpha_c \end{bmatrix}, \tag{S54}$$

with $\alpha_{a,b,c}$ being the polarizability of an ellipsoidal particle along one of its major axis a, b or c:

$$\alpha_a = 3 \cdot V_{ellipsoid} \cdot \frac{n_p^2 - n_m^2}{3 \cdot n_m^2 + 3 \cdot L_a \cdot (n_p^2 - n_m^2)}, \tag{S55}$$

with $V_{ellipsoid}$ being the volume of such an ellipsoid:

$$V_{ellipsoid} = \frac{4}{3}\pi \cdot abc. \tag{S56}$$

The geometric factors $L_{a,b,c}$ are, e.g. for a:

$$L_a = \frac{abc}{2} \cdot \int_0^{+\infty} dv \; \frac{1}{(a^2+v)^{3/2} \cdot (b^2+v)^{1/2} \cdot (c^2+v)^{1/2}}. \tag{S57}$$

Note that the sum of all three geometric factors yields unity:

$$L_a + L_b + L_c = 1. \tag{S58}$$

This means that in the case of a sphere ($a = b = c$), we find that:

$$L_a = L_b = L_c = \frac{1}{3}. \tag{S59}$$

Yielding exactly the result given in Eq. (S21), as then $\alpha_a = \alpha_b = \alpha_c$ and $V_{ellipsoid} \to V_{sphere}$.
When starting to elongate the sphere on one side (here $a > b = c$), we generate a *prolate* spheroid for which the corresponding geometric factor L_a is given as [8]:

$$L_a = \frac{1-e^2}{e^2} \cdot \left(-1 + \frac{1}{2e} \ln \frac{1+e}{1-e} \right), \tag{S60}$$

$$e^2 = 1 - \frac{b^2}{a^2}. \tag{S61}$$

The corresponding geometric factors $L_b = L_c$ can be found using Eq. (S58):

$$L_b = \frac{1}{2} \cdot (1 - L_a). \tag{S62}$$

S8 Deriving the polarizability tensor of a protein from its structure

More interestingly, it is possible to compute the polarizability tensor $\hat{\alpha}$ for a protein, given the knowledge of each atom and its location within the molecule [17, 18] (information which can be obtained from the respective PDB-file of the protein).
The general form of the polarizability tensor $\hat{\alpha}$ is then given by:

$$\hat{\alpha} = \begin{bmatrix} \alpha_{1,1} & \alpha_{1,2} & \alpha_{1,3} \\ \alpha_{2,1} & \alpha_{2,2} & \alpha_{2,3} \\ \alpha_{3,1} & \alpha_{3,2} & \alpha_{3,3} \end{bmatrix}. \tag{S63}$$

The product of polarizability and electric field equates to a certain, induced, dipole moment $\mathbf{p}$:

$$\mathbf{p}_{exc.} = \underbrace{(\hat{\alpha} - \alpha_m)}_{\text{protein - water}} \cdot \underbrace{t_1^{(\perp,\parallel)}(\theta) \cdot \mathbf{E}_{illu}}_{\mathbf{E}_{sample}}, \tag{S64}$$

which we have already written as the *excess* dipole moment $\mathbf{p}_{exc.}$, i.e. including the influence of the surrounding medium. Note that $\mathbf{p}$ is a 3×1 vector, such that:

$$\begin{pmatrix} p^{(1)} \\ p^{(2)} \\ p^{(3)} \end{pmatrix} = \begin{bmatrix} \alpha_{1,1} & \alpha_{1,2} & \alpha_{1,3} \\ \alpha_{2,1} & \alpha_{2,2} & \alpha_{2,3} \\ \alpha_{3,1} & \alpha_{3,2} & \alpha_{3,3} \end{bmatrix} \cdot \begin{pmatrix} E^{(1)}_{sample} \\ E^{(2)}_{sample} \\ E^{(3)}_{sample} \end{pmatrix}. \tag{S65}$$

The total dipole of such a protein $\mathbf{p}$, is assumed to be a superposition of the individual dipole momenta $\mathbf{p}_i$ of all atoms that make up the molecule:

$$\mathbf{p} = \sum_{i=1}^{N} \mathbf{p}_i(\mathbf{E}_{sample}). \tag{S66}$$

According to [18], each of the individual dipole momenta can be expressed as:

$$\mathbf{p}_i(\mathbf{E}_{sample}) = \underbrace{\alpha_i \cdot \mathbf{E}_{sample}}_{\text{atom } i} - \underbrace{\alpha_i \cdot \sum_{j \neq i}^{N} \hat{\mathbf{M}}_{i,j} \cdot \mathbf{p}_j(\mathbf{E}_{sample})}_{\text{all other atoms}}, \tag{S67}$$

with the dipole field tensor $\hat{\mathbf{M}}$ and $\hat{\mathbf{I}}$ being an identity matrix:

$$\hat{\mathbf{M}}_{i,j} = \begin{cases} \frac{1}{\Delta r^3} \cdot \hat{\mathbf{I}} - \frac{3}{\Delta r^5} \begin{bmatrix} \Delta x^2 & \Delta x \Delta y & \Delta x \Delta z \\ \Delta x \Delta y & \Delta y^2 & \Delta y \Delta z \\ \Delta x \Delta z & \Delta y \Delta z & \Delta z^2 \end{bmatrix} & \Delta r > \rho \\ \frac{4\nu^3 - 3\nu^4}{\Delta r^3} \cdot \hat{\mathbf{I}} - \frac{3\nu^4}{\Delta r^5} \begin{bmatrix} \Delta x^2 & \Delta x \Delta y & \Delta x \Delta z \\ \Delta x \Delta y & \Delta y^2 & \Delta y \Delta z \\ \Delta x \Delta z & \Delta y \Delta z & \Delta z^2 \end{bmatrix} & \text{else} \end{cases}, \tag{S68}$$

and $\Delta x = x_i - x_j$ (analogous for Δy & Δz), $\Delta r = \sqrt{\Delta x^2 + \Delta y^2 + \Delta z^2}$, $\nu = \Delta r / \rho$, $\rho = 1.662 \cdot (\alpha_i \alpha_j)^{1/6}$.
To further compute the polarizability of the molecule, we can imagine to sequentially illuminate the protein with a plane wave traveling along the x, y or z-axis, i.e. $\mathbf{E}_l$ with $l \in [x, y, z]$. Yielding a polarizability response in each case as:

$$\hat{\alpha}_l(\mathbf{E}_l) = \begin{pmatrix} \alpha_{l,1} & \alpha_{l,2} & \alpha_{l,3} \end{pmatrix} = \sum_{i=1}^{N} \frac{\mathbf{p}_i(\mathbf{E}_l)}{|\mathbf{E}_l|}. \tag{S69}$$

We then build up the molecular polarizability tensor $\hat{\alpha}$ from these three responses as:

$$\hat{\alpha} = \begin{bmatrix} \hat{\alpha}(\mathbf{E}_x) \\ \hat{\alpha}(\mathbf{E}_y) \\ \hat{\alpha}(\mathbf{E}_z) \end{bmatrix} = \begin{bmatrix} \alpha_{x,1} & \alpha_{x,2} & \alpha_{x,3} \\ \alpha_{y,1} & \alpha_{y,2} & \alpha_{y,3} \\ \alpha_{z,1} & \alpha_{z,2} & \alpha_{z,3} \end{bmatrix}. \tag{S70}$$

The work of [17] has shown that it is possible to perform these computations when evaluating large number of atoms (as typical for proteins) when reformulating it in terms of the following matrix equation:

$$\hat{\mathbf{A}} \cdot \hat{\mathbf{p}} = \hat{\mathbf{E}}. \tag{S71}$$

With $\hat{\mathbf{p}}$ and $\hat{\mathbf{E}}$ being $3N \times 1$ vectors containing the dipole moments and electric fields at each atom ($i \in [1, N]$):

$$\hat{\mathbf{p}} = \begin{pmatrix} \mathbf{p}_1 & \cdots & \mathbf{p}_i & \cdots & \mathbf{p}_N \end{pmatrix}^\top, \tag{S72}$$

$$\hat{\mathbf{E}} = \begin{pmatrix} \mathbf{E}_1 & \cdots & \mathbf{E}_i & \cdots & \mathbf{E}_N \end{pmatrix}^\top, \tag{S73}$$

and each individual component given as a 3×1 vector:

$$\mathbf{p}_i = \begin{pmatrix} p_i^{(1)} & p_i^{(2)} & p_i^{(3)} \end{pmatrix}, \tag{S74}$$

$$\mathbf{E}_i = \begin{pmatrix} E_i^{(1)} & E_i^{(2)} & E_i^{(3)} \end{pmatrix}. \tag{S75}$$

The $3N \times 3N$ matrix $\hat{\mathbf{A}}$ is given as:

$$\hat{\mathbf{A}} = \begin{bmatrix} \hat{\alpha}_{1,1}^{-1} & \hat{\mathbf{M}}_{1,2} & \cdots & \hat{\mathbf{M}}_{1,N} \\ \hat{\mathbf{M}}_{2,1} & \hat{\alpha}_{2,2}^{-1} & \cdots & \hat{\mathbf{M}}_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\mathbf{M}}_{N,1} & \hat{\mathbf{M}}_{N,2} & \cdots & \hat{\alpha}_{N,N}^{-1} \end{bmatrix}, \tag{S76}$$

where each of the elements of $\hat{\mathbf{A}}$ is a 3×3 matrix, e.g. $\hat{\alpha}^{-1}$ is a diagonal matrix such as:

$$\hat{\alpha}_{i,i} = \begin{bmatrix} \alpha_i^{-1} & 0 & 0 \\ 0 & \alpha_i^{-1} & 0 \\ 0 & 0 & \alpha_i^{-1} \end{bmatrix}. \tag{S77}$$

In principle, it is now possible to obtain the dipole moments of each atom according to:

$$\hat{\mathbf{p}} = \hat{\mathbf{A}}^{-1} \cdot \hat{\mathbf{E}}. \tag{S78}$$

With the help of Eq. (S69) and (S70), we can compute the polarizability tensor $\hat{\alpha}$. In practice this is achieved by iteratively solving Eq. (S78) and rewriting $\hat{\mathbf{A}}$ into a sparse matrix through the introduction of a thresholding radius Δr_t (here set to 10 Å), which determines the maximum interaction distance between atoms i and j that we take into account [17].

S9 Mass photometry in the shot-noise limited regime

So far we have only modeled the noise-free signal. Any experiments measuring photon counts are inevitably corrupted by *shot*-noise, which means that the noisy measurement $\mathcal{I}$ is actually given as:

$$\underbrace{\mathcal{I}(\mathbf{r})}_{\substack{\text{noisy} \\ \text{measurement}}} = \underbrace{I(\mathbf{r})}_{\substack{\text{noise-free} \\ \text{expectancy}}} + \underbrace{\mathcal{N}_{\mathcal{I}}(\mathbf{r})}_{\substack{\text{shot} \\ \text{noise}}}, \tag{S79}$$

with $\mathcal{N}_{\mathcal{I}}$ being the individual noise component in a single measurement and I the expectancy $\langle \mathcal{I} \rangle$:

$$\langle \mathcal{I}(\mathbf{r}) \rangle = I(\mathbf{r}). \tag{S80}$$

The noise itself is characterized using a probability distribution connecting the measured outcome $\mathcal{I}$ with the expectancy I:

$$P\left[\mathcal{I}(\mathbf{r}) | I(\mathbf{r})\right] = \frac{I(\mathbf{r})^{\mathcal{I}(\mathbf{r})}}{\mathcal{I}(\mathbf{r})!} \cdot e^{-I(\mathbf{r})}, \tag{S81}$$

which, in case of shot noise, is given as a Poisson distribution with variance being equal to the expectancy [19]:

$$Var\{\mathcal{I}(\mathbf{r})\} = \langle \mathcal{I}(\mathbf{r}) \rangle. \tag{S82}$$

When dealing with large photon counts (as in MP), we can approximate this by a Gaussian distribution:

$$P\left[\mathcal{I}(\mathbf{r}) | I(\mathbf{r})\right] \approx \frac{1}{\sqrt{2\pi \cdot I(\mathbf{r})}} \cdot e^{-\frac{[\mathcal{I}(\mathbf{r}) - I(\mathbf{r})]^2}{2 \cdot I(\mathbf{r})^2}}. \tag{S83}$$

This shot-noise limitation also translates into a noisy representation of the ratiometric contrast:

$$\underbrace{\mathcal{C}(\mathbf{r})}_{\text{noisy}} = \underbrace{C(\mathbf{r})}_{\text{noise-free}} + \underbrace{\mathcal{N}_{\mathcal{C}}(\mathbf{r})}_{\text{noise}}. \tag{S84}$$

Assuming that the Gaussian distribution also holds for the ratiometric signal, we write for the probability distribution:

$$P\left[\mathcal{C}(\mathbf{r}) | C(\mathbf{r}); \sigma_{\mathcal{C}}^2(\mathbf{r})\right] \approx \frac{1}{\sqrt{2\pi \cdot \sigma_{\mathcal{C}}^2(\mathbf{r})}} \cdot e^{-\frac{[\mathcal{C}(\mathbf{r}) - C(\mathbf{r})]^2}{2 \cdot \sigma_{\mathcal{C}}^2(\mathbf{r})}}, \tag{S85}$$

with the measured contrast $\mathcal{C}$, the expectancy C and variance σ_C^2. Note that in case of ratiometric data the variance is *not* equal to the expectancy anymore, as shown in [20]:

$$\langle \mathcal{C}(\mathbf{r}) \rangle = \frac{\langle \mathcal{I}(\mathbf{r}) \rangle}{\langle \mathcal{I}_{bkg}(\mathbf{r}) \rangle} - 1 = C(\mathbf{r}), \tag{S86}$$

$$\sigma_{\mathcal{C}}^2(\mathbf{r}) = Var\{\mathcal{C}(\mathbf{r})\} = \left[\frac{\langle \mathcal{I}(\mathbf{r}) \rangle}{\langle \mathcal{I}_{bkg}(\mathbf{r}) \rangle}\right]^2 \cdot \underbrace{\left[\frac{1}{\langle \mathcal{I}(\mathbf{r}) \rangle} + \frac{1}{\langle \mathcal{I}_{bkg}(\mathbf{r}) \rangle}\right]}_{\approx \frac{2}{\langle \mathcal{I}_{bkg}(\mathbf{r}) \rangle}}, \tag{S87}$$

with $\langle \mathcal{I}_{bkg}(\mathbf{r}) \rangle$ being the expectancy of the measurement without the protein. Note that we assumed that the noise in the raw intensity data is mainly due to the reflected, not the scattered component (true for small proteins). Substituting $\langle \mathcal{C} \rangle$ yields:

$$Var\{\mathcal{C}(\mathbf{r})\} = [\langle \mathcal{C}(\mathbf{r}) \rangle + 1] \cdot \frac{2}{\langle \mathcal{I}_{bkg}(\mathbf{r}) \rangle}. \tag{S88}$$

Since $\langle \mathcal{C}(\mathbf{r}) \rangle \ll 1$ we can express the noise variance as:

$$Var\{\mathcal{C}(\mathbf{r})\} \approx \frac{2}{\langle \mathcal{I}_{bkg}(\mathbf{r}) \rangle}. \tag{S89}$$

Indicating that the noise in the ratiometric distribution is inversely proportional to the expected signal in the raw data $\mathcal{I}$.

S10 Deriving a basic limit on mass sensitivity & resolution in MP

The further analysis assumes no influence from the glass roughness and a scalar representation of the fields involved, i.e. describes the underlying image formation according to:

$$I(\mathbf{r}) = |A_{ref}|^2 + |A_{sca}(\mathbf{r})|^2 + 2 \cdot |A_{ref}| \cdot |A_{sca}(\mathbf{r})|, \tag{S90}$$

with A_{ref} and A_{sca} the electric fields on the detector given by:

$$A_{ref} = E_{ref} \otimes h(\mathbf{r}), \tag{S91}$$

$$A_{sca}(\mathbf{r}) = E_{sca}(\mathbf{r}) \otimes h(\mathbf{r}), \tag{S92}$$

assuming that $h_{ref} = h_{sca} = h$.
We approximate the PSF, corresponding to the intensity signal, as a Gaussian:

$$|h(\mathbf{r})|^2 = h_0^2 \cdot \exp\left(-\frac{|\mathbf{r}|^2}{2 \cdot \sigma_h^2}\right), \tag{S93}$$

with $\sigma_h = 0.21 \cdot \lambda / NA$ given in a least-squares sense [21]. The amplitude h_0^2 of this Gaussian has the physical meaning of an irradiance (units: W/m^2) and can be related to a power P_0 (units: W) distributed over the *BFP* area A_{BFP} as stated in [22]:

$$h_0^2 = \frac{P_0 \cdot A_{BFP}}{\lambda^2 \cdot f_{obj}^2}. \tag{S94}$$

The *BFP* area is simply given by integrating a constant field distribution over polar coordinates ρ and ϕ, with $\rho = f_{obj} \cdot n_i \cdot \sin\theta$.

$$A_{BFP} = \int_0^{\rho_c} d\rho \int_0^{2\pi} d\phi \;\; \rho = \pi \cdot f_{obj}^2 \;\; \cdot \;\; \left(n_i \cdot \frac{n_m}{n_g}\right)^2, \tag{S95}$$

where ρ_c is the radial distance corresponding to the critical angle $\theta_c = \arcsin(n_m/n_g)$ of the glass - water interface and f_{obj} the focal length of the detection objective, given as the ratio of tube lens focal length to objective magnification: $f_{obj} = f_{tube}/M_{obj}$. We further modify this *BFP*-area to include the *aplanatic* factor [2] by computing:

$$A_{BFP} = \int_0^{\rho_c} d\rho \int_0^{2\pi} d\phi \;\; \frac{\rho}{\cos\theta(\rho)} = \int_0^{\rho_c} d\rho \int_0^{2\pi} d\phi \;\; \frac{\rho}{\cos\left[\arcsin\left(\frac{\rho}{f_{obj} \cdot n_i}\right)\right]} = \pi \cdot f_{obj}^2 \;\; \cdot \;\; 2 \cdot n_i^2 \cdot \left[1 - \sqrt{1 - \left(\frac{n_m}{n_g}\right)^2}\right]. \tag{S96}$$

This leads to an effective increase of A_{BFP}, which we indicate by the enhancement factor γ:

$$A_{BFP} \quad \rightarrow \quad \gamma \cdot A_{BFP}, \tag{S97}$$

with $\gamma = 1.36$ for a 1.42 *NA* oil-immersion objective. Additionally we include the effects of the near-field contribution of the scatterer on the glass-water interface by azimuthally (numerically) averaging Eq. (S26) & Eq. (S27), comparing the resulting (integrated) field distribution with that of Eq. (S95) and find that $\gamma \approx 1.58$ for the same 1.42 *NA* objective.

NA	1.2	1.3	1.4	1.5
n_i	1.515	1.515	1.515	1.515
γ	1.14	1.21	1.53	1.67

Table S1. Enhancement factor γ for different detection objective *NA*s.

In terms of the scattering, we write the electric field as a delta-distribution:

$$E_{sca}(\mathbf{r}) = t_2 \cdot t_1 \cdot \delta(\mathbf{r}). \tag{S98}$$

The convolution $\delta \otimes h = h$ yields the PSF itself, hence the scattered field on the detector is given as:

$$\begin{aligned} A_{sca}(\mathbf{r}) &= E_{sca}(\mathbf{r}) \otimes h(\mathbf{r}) = \\ &= t_2 \cdot t_1 \cdot \int_{-\infty}^{+\infty} d\mathbf{v}\ \delta(\mathbf{v}-\mathbf{r}) \cdot h(\mathbf{v}) = t_2 \cdot t_1 \cdot h(\mathbf{r}). \end{aligned} \tag{S99}$$

The scattered power (units W) of the protein is given by:

$$P_{sca} = \mu \cdot \sigma_{sca} \cdot I_{illu}, \tag{S100}$$

with μ the collection efficiency of the detection objective, I_{illu} the illumination intensity (units: W/m^2) and the scattering cross-section (units: m^2), given as:

$$\sigma_{sca} = \frac{1}{6\pi} \cdot \left(\frac{2\pi}{\lambda} \cdot n_m \right)^4 \cdot \alpha^2. \tag{S101}$$

The polarizability α (units: m^3) of a spherical particles is given as:

$$\alpha = 3 \cdot V \cdot \frac{n_p^2 - n_m^2}{n_p^2 + 2 \cdot n_m^2}, \tag{S102}$$

with V being the volume (units: m^3) and n_p, n_m the refractive indices of protein and surrounding medium.
Here, however, we want to simply state a linear relationship between the proteins mass m and α:

$$\alpha = \delta\alpha \cdot m, \tag{S103}$$

$$\delta\alpha = 723.857\ \frac{\text{Å}^3}{\text{kDa}}, \tag{S104}$$

with $\delta\alpha$ which is the slope of the fitted line in Figure 4a in the main text.
In terms of detected, *scattered*, intensity this yields (note that t_2 is already taken care of in γ):

$$\begin{aligned} |A_{sca}(\mathbf{r})|^2 &= |t_1|^2 \cdot h_0^2 \cdot \exp\left(-\frac{|\mathbf{r}|^2}{2 \cdot \sigma_h^2} \right) \\ &= \underbrace{|t_1|^2 \cdot \frac{\mu \cdot \sigma_{sca} \cdot I_{illu} \cdot \gamma \cdot A_{BFP}}{\lambda^2 \cdot f_{obj}^2}}_{|A_{sca}|_0^2} \cdot \exp\left(-\frac{|\mathbf{r}|^2}{2 \cdot \sigma_h^2} \right) \\ &= |A_{sca}|_0^2 \cdot \exp\left(-\frac{|\mathbf{r}|^2}{2 \cdot \sigma_h^2} \right). \end{aligned} \tag{S105}$$

Since the illumination is assumed to be widefield, the reflection simply takes the form of a *constant* field distribution:

$$E_{ref} = \tau \cdot r. \tag{S106}$$

With this we are able to express the detected field A_{ref} as:

$$|A_{ref}|^2 = \underbrace{|\tau|^2 \cdot |r|^2 \cdot I_{illu}}_{=|A_{ref}|_0^2} = |A_{ref}|_0^2. \tag{S107}$$

Note that both, A_{ref} and A_{sca}, are in SI base units of W/m^2.
To compute the effect of shot noise, we need to converte the two fields, scattered & reflected, into photon counts. For this we recall the definition of power as the ratio of energy change to time duration required for this change to happen:

$$P = \frac{\Delta E}{\Delta t} \qquad \rightarrow \qquad \Delta E = P \cdot \Delta t. \tag{S108}$$

In our case Δt correspond to the (effective) exposure time when recording light on our detector.
The energy of a single photon at wavelength λ is:

$$E_{ph} = h \cdot \frac{c}{\lambda}. \tag{S109}$$

With this we are able to compute the number of photons per (light) power (units: W), according to:

$$N = \frac{\Delta E}{E_{ph}} = \frac{P \cdot \Delta t \cdot \lambda}{h \cdot c}. \tag{S110}$$

Hence, the number of detected, *reflected*, photons per pixel is now given as:

$$\begin{aligned} N_{ref} &= OT \cdot QE \cdot \left[\int_{-d_{px}/2}^{+d_{px}/2} dx \int_{-d_{px}/2}^{+d_{px}/2} dy \ |A_{ref}|^2 \right] \cdot \frac{\Delta t \cdot \lambda}{h \cdot c} \\ &= OT \cdot QE \cdot |A_{ref}|_0^2 \cdot \left[\int_{-d_{px}/2}^{+d_{px}/2} dx \int_{-d_{px}/2}^{+d_{px}/2} dy \ 1 \right] \cdot \frac{\Delta t \cdot \lambda}{h \cdot c} \\ &= OT \cdot QE \cdot |A_{ref}|_0^2 \cdot A_{px} \cdot \frac{\Delta t \cdot \lambda}{h \cdot c}. \end{aligned} \tag{S111}$$

with the integration indicating the effect of the photon-sensitive area of a single pixel $A_{px} = d_{px}^2$, *OT* being the throughput of the optical system and the quantum efficiency *QE* of the detector.
Similarly we find for the number of detected, *scattered*, photons:

$$\begin{aligned} N_{sca} &= OT \cdot QE \cdot \left[\int_{-d_{px}/2}^{+d_{px}/2} dx \int_{-d_{px}/2}^{+d_{px}/2} dy \ |A_{sca}|^2 \right] \cdot \frac{\Delta t \cdot \lambda}{h \cdot c} \\ &= OT \cdot QE \cdot |A_{sca}|_0^2 \cdot \left[\int_{-d_{px}/2}^{+d_{px}/2} dx \int_{-d_{px}/2}^{+d_{px}/2} dy \ \exp\left(-\frac{x^2+y^2}{2 \cdot \sigma_h^2} \right) \right] \cdot \frac{\Delta t \cdot \lambda}{h \cdot c} \\ &= OT \cdot QE \cdot |A_{sca}|_0^2 \cdot \underbrace{\left[\sqrt{2\pi} \cdot \sigma_h \cdot \mathrm{erf}\left(\frac{d_{px}}{2\sqrt{2} \cdot \sigma_h} \right) \right]^2}_{\approx A_{px}} \cdot \frac{\Delta t \cdot \lambda}{h \cdot c}. \end{aligned} \tag{S112}$$

This allows us now to define the photon count with (N) and without (N_{bkg}) scatterer:

$$N = N_{ref} + N_{sca} + 2 \cdot \sqrt{N_{ref} \cdot N_{sca}}, \tag{S113}$$

$$N_{bkg} = N_{ref}. \tag{S114}$$

In the following we are assuming that $N \leq N_{FWD}$, the full-well-depth of the detector. Note that one effectively increases the detected signal beyond N_{FWD} by employing pixel and frame binning (denoted by n_{bin}):

$$N = n_{avg} \cdot n_{bin} \cdot \left(N_{ref} + N_{sca} + 2 \cdot \sqrt{N_{ref} \cdot N_{sca}} \right), \tag{S115}$$

$$N_{bkg} = n_{avg} \cdot n_{bin} \cdot N_{ref}, \tag{S116}$$

with n_{avg} being the window size employed in the moving average computation of the ratiometric signal.
The ratiometric contrast in terms of photon counts C_N is:

$$\langle \mathcal{C}_\mathcal{N} \rangle = C_N = \frac{N - N_{bkg}}{N_{bkg}}. \tag{S117}$$

At the same time the variance of this ratiometric signal approximately given as (see Eq. (S89)):

$$Var\{\mathcal{C}_\mathcal{N}\} \approx \frac{2}{N_{bkg}}. \tag{S118}$$

Hence, we can define the signal-to-noise ratio *SNR* of the ratiometric signal as:

$$SNR = \frac{\langle \mathcal{C}_\mathcal{N} \rangle}{\sqrt{Var\{\mathcal{C}_\mathcal{N}\}}} = \frac{1}{\sqrt{2}} \cdot C_N \cdot \sqrt{N_{bkg}} = \frac{1}{\sqrt{2}} \cdot \frac{N - N_{bkg}}{\sqrt{N_{bkg}}}. \tag{S119}$$

Altogether, the number of detected (*reflected*) photons is given as:

$$N_{bkg} = N_{ref} = I_{illu} \cdot \frac{\Delta t \cdot \lambda}{h \cdot c} \cdot OT \cdot QE \cdot n_{avg} \cdot n_{bin} \cdot |\tau|^2 \cdot |r|^2 \cdot d_{px}^2. \tag{S120}$$

Similarly, we find for the scattered component ($|t_1|^2 = 1 - |r|^2$):

$$N_{sca} = I_{illu} \cdot \frac{\Delta t \cdot \lambda}{h \cdot c} \cdot \underbrace{\frac{(2\pi)^4}{6\pi}}_{=\frac{8}{3}\pi^3} \cdot n_{avg} \cdot n_{bin} \cdot \frac{OT \cdot QE \cdot \left[1 - |r|^2\right] \cdot \mu \cdot n_m^4 \cdot \gamma \cdot A_{BFP} \cdot d_{px}^2}{\lambda^6 \cdot f_{obj}^2} \cdot \left[\delta\alpha \cdot m\right]^2. \tag{S121}$$

The interference component is now defined as:

$$2 \cdot \sqrt{N_{ref} \cdot N_{sca}} = I_{illu} \cdot \frac{\Delta t \cdot \lambda}{h \cdot c} \cdot 2\sqrt{\frac{8}{3}\pi^3} \cdot n_{avg} \cdot n_{bin} \cdot \frac{OT \cdot QE \cdot \sqrt{1 - |r|^2} \cdot \sqrt{\mu} \cdot n_m^2 \cdot \sqrt{\gamma} \cdot \sqrt{A_{BFP}} \cdot d_{px}^2 \cdot |\tau| \cdot |r|}{\lambda^3 \cdot f_{obj}} \cdot \delta\alpha \cdot m. \tag{S122}$$

Hence, the differential signal, neglecting the purely scattering component, is given as:

$$\begin{aligned} N - N_{bkg} &\approx 2 \cdot \sqrt{N_{ref} \cdot N_{sca}} = \\ &= I_{illu} \cdot \frac{\Delta t \cdot \lambda}{h \cdot c} \cdot 2\sqrt{\frac{8}{3}\pi^3} \cdot n_{avg} \cdot n_{bin} \cdot \frac{OT \cdot QE \cdot \sqrt{1 - |r|^2} \cdot \sqrt{\mu} \cdot n_m^2 \cdot \sqrt{\gamma} \cdot \sqrt{A_{BFP}} \cdot d_{px}^2 \cdot |\tau| \cdot |r|}{\lambda^3 \cdot f_{obj}} \cdot \delta\alpha \cdot m. \end{aligned} \tag{S123}$$

And the square-root of the reflected component as:

$$\sqrt{N_{bkg}} = \sqrt{I_{illu}} \cdot \sqrt{\frac{\Delta t \cdot \lambda}{h \cdot c}} \cdot \sqrt{n_{avg} \cdot n_{bin}} \cdot \sqrt{OT \cdot QE} \cdot |\tau| \cdot |r| \cdot d_{px}. \tag{S124}$$

Finally we are able to write for the ratiometric SNR:

$$\begin{aligned} SNR &= \frac{1}{\sqrt{2}} \cdot \frac{N - N_{bkg}}{\sqrt{N_{bkg}}} \approx \frac{2}{\sqrt{2}} \cdot \sqrt{N_{sca}} = \\ &= \sqrt{I_{illu}} \cdot \sqrt{\frac{\Delta t \cdot \lambda}{h \cdot c}} \cdot \frac{2}{\sqrt{2}} \sqrt{\frac{8}{3}\pi^3} \cdot \sqrt{n_{avg} \cdot n_{bin}} \cdot \frac{\sqrt{OT \cdot QE} \cdot \sqrt{1 - |r|^2} \cdot \sqrt{\mu} \cdot n_m^2 \cdot \sqrt{\gamma} \cdot \sqrt{A_{BFP}} \cdot d_{px}}{\lambda^3 \cdot f_{obj}} \cdot \delta\alpha \cdot m. \end{aligned} \tag{S125}$$

Note that the *SNR* is independent of the mask strength τ, as it influences contrast and noise simultaneously, such that they cancel each other. Contrary to this, the ratiometric contrast is improved upon modifying the effective reflectivity from the glass-buffer medium interface:

$$\langle \mathcal{C}_\mathcal{N} \rangle = 2\sqrt{\frac{8}{3}\pi^3} \cdot \frac{\sqrt{1 - |r|^2} \cdot \sqrt{\mu} \cdot n_m^2 \cdot \sqrt{\gamma} \cdot \sqrt{A_{BFP}}}{\lambda^3 \cdot f_{obj} \cdot |\tau| \cdot |r|} \cdot \delta\alpha \cdot m. \tag{S126}$$

But the contrast is not simply enhanced by increasing I_{illu}, n_{avg}, n_{bin} or any other parameter that yields a higher photon flux at the detector. Those merely affect the noise of the ratiometric signal, as:

$$Var\{\mathcal{C}_\mathcal{N}\} \approx \frac{2}{I_{illu} \cdot \frac{\Delta t \cdot \lambda}{h \cdot c} \cdot OT \cdot QE \cdot n_{avg} \cdot n_{bin} \cdot |\tau|^2 \cdot |r|^2 \cdot d_{px}^2}. \tag{S127}$$

We can use the analytical model to compute a *mass-equivalent* signal-to-noise ratio: the smallest mass m_q that one would be able to detect with an SNR-level of q:

$$m_q = \frac{q}{\delta\alpha}\cdot\frac{1}{\sqrt{I_{illu}}}\cdot\sqrt{\frac{h\cdot c}{\Delta t\cdot\lambda}}\cdot\frac{\sqrt{2}}{2}\sqrt{\frac{3}{8\pi^3}}\cdot\frac{1}{\sqrt{n_{avg}\cdot n_{bin}}}\cdot\frac{\lambda^3\cdot f_{obj}}{\sqrt{OT\cdot QE}\cdot\sqrt{1-|r|^2}\cdot\sqrt{\mu}\cdot n_m^2\cdot\sqrt{\gamma}\cdot\sqrt{A_{BFP}}\cdot d_{px}}. \tag{S128}$$

In terms of mass *resolution* we employ the concept of the Quantum-Cramer-Rao lower bound (QCRLB) [23]. The smallest achievable uncertainty of estimating the mass of the protein is then given as:

$$\begin{aligned}\sigma_m &= \frac{1}{2}\cdot\frac{m}{\sqrt{N_{sca}}} = \frac{m}{\sqrt{2}}\cdot\frac{1}{SNR} = m_{q=\sqrt{0.5}} = \\ &= \frac{1}{2}\cdot\frac{1}{\delta\alpha}\cdot\frac{1}{\sqrt{I_{illu}}}\cdot\sqrt{\frac{h\cdot c}{\Delta t\cdot\lambda}}\cdot\sqrt{\frac{3}{8\pi^3}}\cdot\frac{1}{\sqrt{n_{avg}\cdot n_{bin}}}\cdot\frac{\lambda^3\cdot f_{obj}}{\sqrt{OT\cdot QE}\cdot\sqrt{1-|r|^2]}\cdot\sqrt{\mu}\cdot n_m^2\cdot\sqrt{\gamma}\cdot\sqrt{A_{BFP}}\cdot d_{px}}.\end{aligned} \tag{S129}$$

Note that this uncertainty independent on whether the measurement is performed interferometrically or in darkfield mode. The latter, however, exhibits a worsened SNR by a factor of $\sqrt{2}$, as:

$$SNR_{darkfield} = \frac{N_{sca}}{\sqrt{N_{sca}}} = \sqrt{N_{sca}} = \frac{1}{\sqrt{2}}\cdot SNR, \tag{S130}$$

with:

$$\sqrt{N_{sca}} = \sqrt{I_{illu}}\cdot\sqrt{\frac{\Delta t\cdot\lambda}{h\cdot c}}\cdot\sqrt{\tfrac{8}{3}\pi^3}\cdot\sqrt{n_{avg}\cdot n_{bin}}\cdot\frac{\sqrt{OT\cdot QE}\cdot\sqrt{1-|r|^2}\cdot\sqrt{\mu}\cdot n_m^2\cdot\sqrt{\gamma}\cdot\sqrt{A_{BFP}}\cdot d_{px}}{\lambda^3\cdot f_{obj}}\cdot\delta\alpha\cdot m. \tag{S131}$$

Making it more difficult to reach the fundamental σ_m-value in practice. Of course this also translates into the smallest detectable mass: $m_{q,darkfield} = \sqrt{2}\cdot m_q$.

S11 Signal-to-noise ratio including excess noise

To including the additional excess (or baseline) noise we add the additional uncertainty $\sigma_{exc.}^2$ (in terms of the ratiometric signal) to the variance stemming from shot-noise alone, according to:

$$SNR_{exc.} = \frac{\langle\mathcal{C}_\mathcal{N}\rangle}{\sqrt{Var\{\mathcal{C}_\mathcal{N}\}+\sigma_{exc.}^2}} = \frac{C_N}{\sqrt{\frac{2}{N_{bkg}}+\sigma_{exc.}^2}}. \tag{S132}$$

We can rearrange the denominator which yields in:

$$\sqrt{\frac{2}{N_{bkg}}+\sigma_{exc.}^2} = \sqrt{\frac{2 + \sigma_{exc.}^2\cdot N_{bkg}}{N_{bkg}}}. \tag{S133}$$

Altogether we find for the signal-to-noise ratio including the excess noise:

$$SNR_{exc.} = \frac{N-N_{bkg}}{N_{bkg}}\cdot\sqrt{N_{bkg}}\cdot\frac{1}{\sqrt{2+\sigma_{exc.}^2\cdot N_{bkg}}} = \tag{S134}$$

$$= \frac{N-N_{bkg}}{\sqrt{N_{bkg}}}\cdot\frac{1}{\sqrt{2+\sigma_{exc.}^2\cdot N_{bkg}}}. \tag{S135}$$

Which we can relate to the shot-noise limited *SNR* through:

$$SNR_{exc.} = SNR\cdot\underbrace{\frac{\sqrt{2}}{\sqrt{2+\sigma_{exc.}^2\cdot N_{bkg}}}}_{=\xi}. \tag{S136}$$

Verifying that for $\sigma_{exc.} = 0$ we get: $SNR_{exc.} = SNR$. Note that this automatically means that the $SNR_{exc.}$ equivalent mass $m_{q,exc.}$ is still given by Eq. (S128), only with the substitution:

$$q \;\rightarrow\; q/\xi. \tag{S137}$$

Finally we would want to see how $SNR_{exc.}$ behaves when detecting a larger number of photons, e.g. through increase of the illumination (in case of pure shot-noise the SNR is unlimited as: $SNR^{max} = +\infty$). For this we introduce a scaling factor ϵ:

$$N \rightarrow \epsilon \cdot \left(N_{bkg} + N_{sca} + 2 \cdot \sqrt{N_{bkg} \cdot N_{sca}} \right), \tag{S138}$$

$$N_{bkg} \rightarrow \epsilon \cdot N_{bkg}. \tag{S139}$$

So we get for the shot-noise limited *SNR*:

$$\sqrt{2} \cdot SNR \approx \frac{\epsilon \cdot 2 \cdot \sqrt{N_{bkg} \cdot N_{sca}}}{\sqrt{\epsilon \cdot N_{bkg}}} = \sqrt{\epsilon} \cdot 2 \cdot \sqrt{N_{sca}}. \tag{S140}$$

With this we find for $SNR_{exc.}$:

$$SNR_{exc.} = 2 \cdot \sqrt{\frac{\epsilon \cdot N_{sca}}{2 + \sigma_{exc.}^2 \cdot \epsilon \cdot N_{bkg}}} = 2 \cdot \sqrt{\frac{N_{sca}}{\frac{2}{\epsilon} + \sigma_{exc.}^2 \cdot N_{bkg}}}. \tag{S141}$$

Taking this in the limit of $\epsilon \rightarrow \infty$:

$$\lim_{\epsilon \rightarrow \infty} SNR_{exc.} = \frac{2}{\sigma_{exc.}} \cdot \sqrt{\frac{N_{sca}}{N_{bkg}}}. \tag{S142}$$

With the following ratio of scattered to reflected number of photons:

$$\begin{aligned} \frac{N_{sca}}{N_{bkg}} &= \frac{|A_{sca}|_0^2}{|A_{ref}|_0^2} = \frac{|t_1|^2 \cdot \frac{\mu \cdot \sigma_{sca} \cdot \gamma \cdot A_{BFP}}{\lambda^2 \cdot f_{obj}^2}}{|\tau|^2 \cdot |r|^2} = \\ &= \frac{8}{3} \cdot \pi^3 \cdot \frac{|t_1|^2 \cdot \mu \cdot n_m^4 \cdot \gamma \cdot A_{BFP}}{\lambda^6 \cdot f_{obj}^2 \cdot |\tau|^2 \cdot |r|^2} \cdot [\delta\alpha \cdot m]^2 . \end{aligned} \tag{S143}$$

S12 Mass photometry signal for different mask types

Figure S1 presents line profiles of the simulated data for Hsp16.5 (24-mer; ≈ 400 kDa) in the lateral (a & c) and axial (b) dimensions. In general increasing the mask strength increases the sensitivity of the MP-system as the small ratiometric contrast (a & b) gets enhanced, such that the signal rises above the shot-noise level. To obtain the optimum performance it is necessary to further make sure that the reference and scattered light are exactly in-phase, which requires the mask to introduce an additional $\pi/2$ phase shift. We term this the *phase matched* mask (a & b, bottom) and now observe the optimum contrast and the nominal focal plane. Note that, when considering a finite full-well-depth of the detector, the attenuation mask can not be arbitrarily strong. As the recorded modulation due to glass roughness increases such, that certain pixels will be saturated.

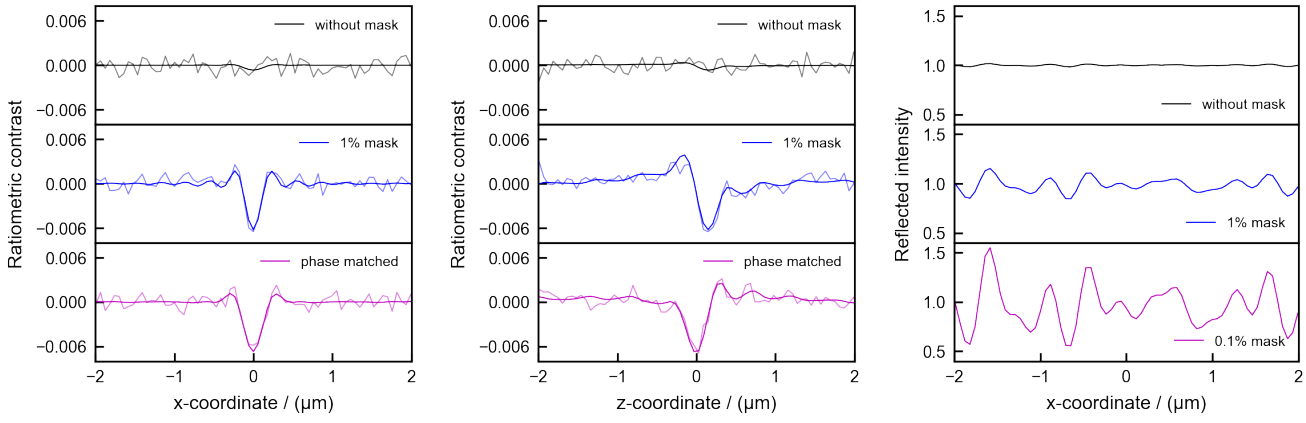
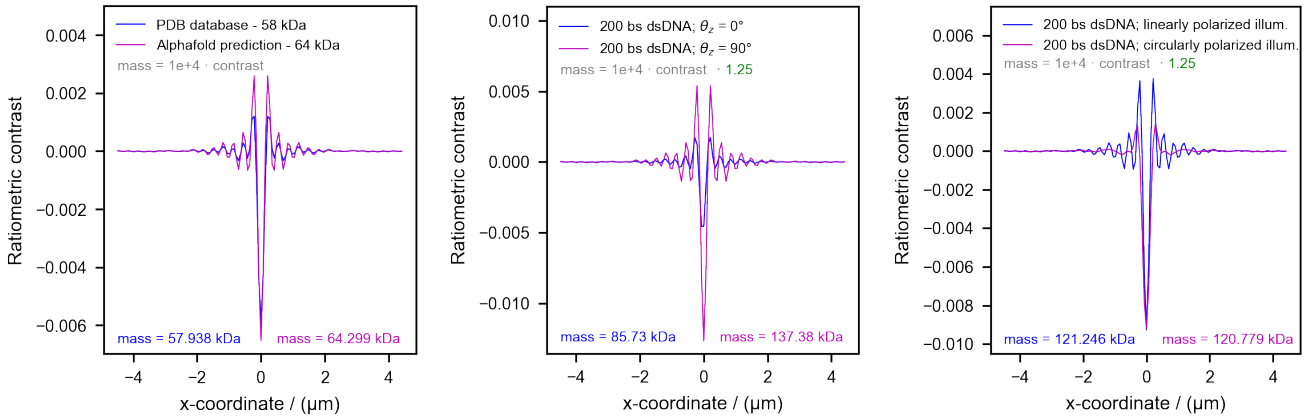


Figure S1. Ratiometric contrast and recorded intensity when simulating the MP signal of Hsp16.5 (24-mer, ≈ 400 kDa). a) Lateral line profile of the ratiometric contrast, when imaging using no mask (top, black), a 1% mask (middle, blue) and a 1% mask with phase matching condition (bottom, magenta). The thick line represents the expected (noise-free) signal and the thin line a noisy measurement. (b) The same as in (a), only as a axial line profile. Note how the axial position of optimum contrast is shifted towards the nominal focal plane, when using the phase matched 1% mask. (c) Lateral line profile of the recorded intensity, including the effects due to glass roughness. Improving the performance of the MP-system by increasing the mask strength is limited, as this leads to a stronger detected modulation in the raw data, which ultimately fills the complete full-well-depth of the detector, i.e. preventing more information to be captured.

S13 Comparison between PDB vs alphafold structure of BSA

As noted in the main text, the PDB-file that we used to simulate the ratiometric signal of BSA (PDBID 3V03) does only contain a smaller subset of the atoms of a real BSA molecule in buffer solution. To get a more realistic estimation of the inferred contrast we performed the same simulation with the polarizability obtained from a PDB-structured predicted by AlphaFold [24] (UniProt P02769). Overall we observe an increase of $\approx 11\%$, resulting in $m \approx 64$ kDa, in good agreement with the experimentally expected values.

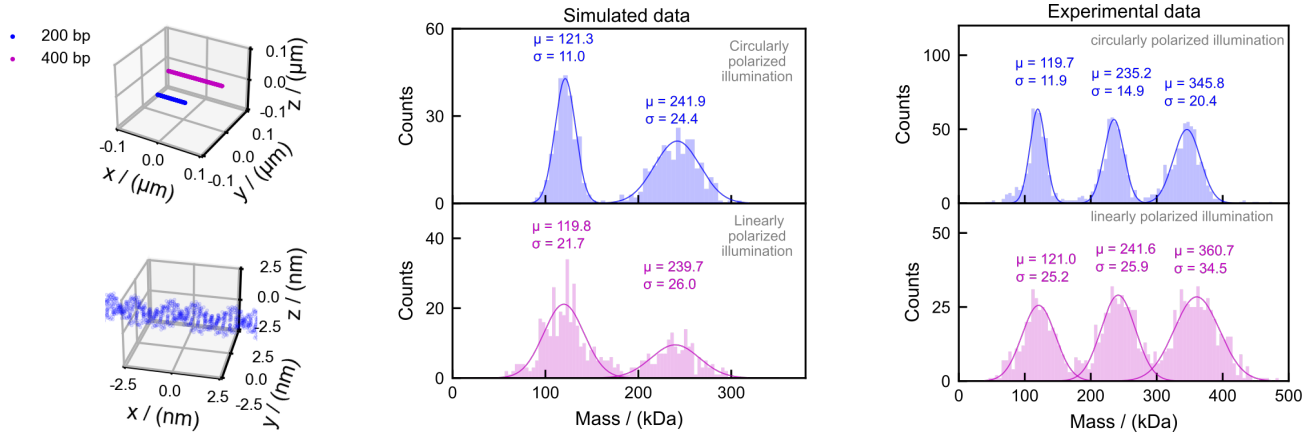


(a) (b) (c)

Figure S2. Ratiometric contrast comparison between PDB & alphafold prediction of BSA (a), 200 bp double-stranded (ds) DNA at different orientations (b) and illumination polarizations. a) The AlphaFold prediction of BSA (UniProt P02769) contains the atoms that were missing in the original PDB structure (PDBID 3V03), yielding an overall increase of the inferred mass by $\approx 11\%$. b) When imaging 200 bp dsDNA using linearly polarized illumination the ratiometric signal changes with the orientation ($\theta_z = 0°$, $\theta_z = 90°$) of the rod-like molecule. c) Ratiometric signal for randomly oriented dsDNA molecule imaged with linearly (blue) or circularly (magenta) polarized light, yielding the same contrast which corresponds to ≈ 121 kDa. Note that the mass-to-contrast conversion is larger when imaging dsDNA, compared to proteins (see [25]).

S14 MP signal for dsDNA imaged with linearly & circularly polarized illumination

In addition to the data shown in Fig3 of the main text, here we include measurements for 600 bp DNA. We further investigated the theoretical signal to be expected from such a rod-like structure. For this we first generated a random sequence of dsDNA containing 200 bp which we transferred into a PDB-file (B-DNA model) using [26] (see blue in Fig. S3 a). The simulated ratiometric signal for two orthogonal orientations ($\theta_z = 0°$, $\theta_z = 90°$) are shown in Fig. S2 b), indicating a strong change depending how the rod-like molecule is aligned with respect to the linearly polarized illumination. We also investigate the effect of randomly oriented dsDNA (200 bp) when being imaged either with linearly (blue) or circularly (magenta) polarized light. Both imaging scenarios yield the same ratiometric contrast amounting to ≈ 121 kDa. Note the larger mass-to-contrast conversion factor ($\times 1.25$) when imaging dsDNA compared to that of proteins, as reported in [25].



(a) (b) (c)

Figure S3. a) Graphical depiction of the used DNA model. Top: Double-stranded (ds) DNA PDB-model with 200 (blue) and 400 (magenta) base pairs, showing the assumed rod-like shape ($\varnothing \approx 2$ nm; length ≈ 68 & 136 nm). Bottom: zoom depicting the helical structure of the 200 bp dsDNA. b) Simulated landing assay histogram of dsDNA (200 bp: $m \approx 120$ kDa; 400 bp: $m \approx 240$ kDa), imaged using circularly (top, blue) and linearly (bottom, magenta) polarized light. Indicating the broadening of mass peaks, depending on the illumination polarization. c) Experimental mass histogram of a mixture of 200, 400 and 600 bp dsDNA.

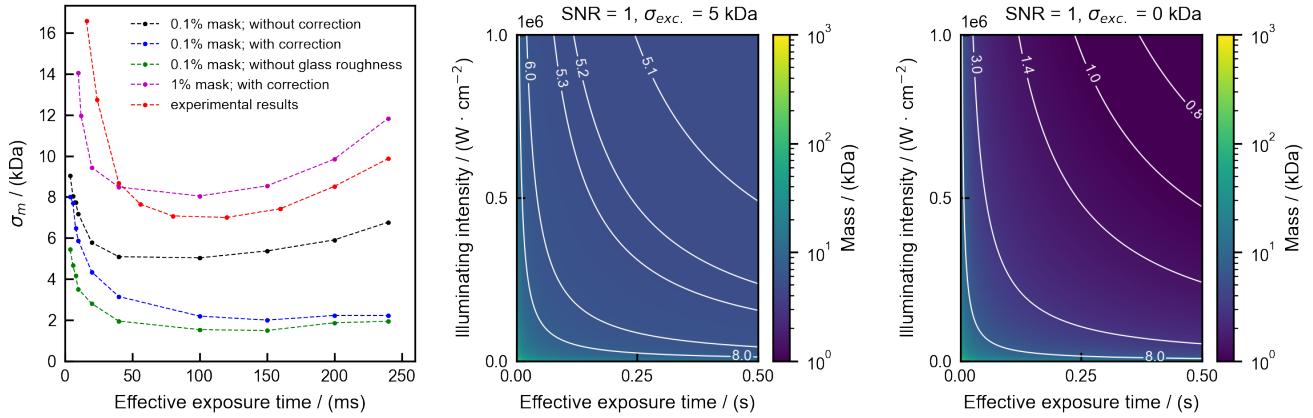
Further, we simulated landing assays with the 200 & 400 bp version of the dsDNA, and depict the fitted mass histograms in a

similar way to the experimental findings (Fig. 3a) in Figure S3 b). In case of the 200 bp dsDNA we see the same broadening of the mass peak, depending on the illumination polarization. For the 400 bp version, however, our simulation indicates a much broader peak already when using circularly polarized light. This is mainly due to the fact that the simulation only takes into account point-like objects, i.e., particles smaller than the simulated pixel size (here 70 nm). In the presented case, this causes an error leading to the apparent mass broadening, as the 400 bp dsDNA molecules have a length of ≈ 136 nm, while the 200 bp dsDNA only amounts to ≈ 68 nm. Meaning that in the case of the 400 bp dsDNA, the pixelation error hides the actual broadening due to the linearly vs circularly polarized illumination (as experimentally shown in Fig. 3a in the main text).

Note that increasing the pixel size will not alleviate the error caused by the pixelation, as it would still be required to fulfil the Shannon-Nyquist sampling theorem [27], i.e. a pixel size of $\lesssim 80$ nm when using a numerical aperture of $NA = 1.42$ and $\lambda = 445$ nm. A potential approach to circumvent the pixelation error, would be to modify the simulation routine such that the convolution in Eq. (S1) is performed on a superset of pixels that encode the orientation of the DNA molecule. However, this has been beyond the scope of our work.

S15 Simulated mass resolution for detecting BSA monomer at different exposure times

Being able to simulate landing assay movies we looked into the mass resolution, i.e. the width of the fitted Gaussian, when detecting the BSA-monomer alone. We do this for a changing effective exposure time and assume shot-noise limited performance. The results are shown in Fig. S4 a) for different simulated scenarios and experimental results.
Starting out with a 0.1% mask and simulated glass roughness without the correction in post-processing (black curve) we observe an optimum value of $\sigma_m \approx 5.5$ kDa. Lowering the effective exposure time will make the analyzed ratiometric movies look more noisy, making it more difficult for the particle picking & fitting algorithms to perform optimally. Making use of the correction step (Eq. (S51); blue) enables to extract a minimum $\sigma_m \approx 2$ kDa, in this ideal simulated scenario. Interestingly, performing the simulation without the glass roughness (hence no need for such a correction; green), yields a similar value for σ_m. Indicating that in principle the correction is enabling to achieve a performance similar to that of using a perfectly flat coverslip, assuming a constant $|\mathbf{E}_{illu}|$. Experimental results are shown in red, exhibiting a minimum $\sigma_m \approx 7.5$ kDa, which fits more to a 1% mask. This difference is most likely due to the additional baseline noise of ≈ 5 kDa (see Fig. 5 in main text).



(a) **(b)** **(c)**

Figure S4. a) Fitted width of Gaussian for simulated and experimental landing assay data of the BSA monomer, in dependency of effective exposure time. Simulating a 0.1% mask including glass roughness yields a minimum $\sigma \approx 5.5$ kDa, while this can be further decreased when making use of the correction step in post-processing to $\sigma \approx 2$ kDa. Note that this comes very close to the value we observe when simulating without glass roughness. Experimental results, however, achieve $\sigma \approx 7.5$ kDa, which most likely is due to the additional baseline noise of ≈ 5 kDa as described in the main text (e.g. see Fig. 5 b - c) Minimum detectable mass with $SNR = 1$ with (b) and without (c) this additional excess noise. For realistic imaging parameters we find that $m_{q=1} \approx 5$ kDa, which is in agreement with the results shown in [28].

S16 Smallest detectable mass with $SNR = 1$; with and without excess noise

Equation Eq. (S128) enables us to calculate the smallest detectable mass given a certain *SNR* level. Here we show the results for the fundamental limit, i.e. $SNR = 1$. When including the additional baseline noise of ≈ 5 kDa and realistic imaging parameters, we observe $m_{q=1} \approx 5$ kDa. Which seems realistic given the results shown in [28], detecting a 9 kDa protein at $SNR \approx 1.4$. When neglecting the additional baseline noise our calculation suggest to reach $m_{q=1} \approx 1.5$ kDa with the same illumination power & effective integration time.

S17 Influence of the reflection correction on particle number and their landing coordinates

To compensate for the spatially varying reflection due to the glass roughness, we employ a correction step (Eq. (S51))[16], which alters the ratiometric contrast C locally. This might affect the subsequent data analysis, especially the particle picking which depends on the local distribution of the corrected C'. To get an estimate on these potential alterations we analyzed an experimental data set, with and without employing eq. S51, and report the results in Table S2 & S3 below.

	Number of *found* particles	Number of *fitted* particles
Without reflection correction	1886	1348
With reflection correction	1870	1344
Ratio of overall events	∼99%	∼99%

Table S2. Number of found/fitted particles when processing the raw data with and without reflection correction.

	Found particles	*Fitted* particles
Number of *matching* landing coordinates [x,y,frame]	1710	1225
Ratio to overall detected particles	∼91%	∼91%

Table S3. Number of matching landing coordinates when processing the raw data with and without reflection correction.

We see that the correction step maintains almost all landing events ($\sim 99\%$), while not changing the found/fitted landing coordinates ($\sim 91\%$ matching in x,y and frame number).

References

[1] Joseph W Goodman and P Sutton. Introduction to fourier optics. *Quantum and Semiclassical Optics-Journal of the European Optical Society Part B*, 8(5):1095, 1996.

[2] Lukas Novotny and Bert Hecht. *Principles of nano-optics*. Cambridge university press, 2012.

[3] Ratsimandresey Holinirina Dina Miora, Erich Rohwer, Martin Kielhorn, Colin JR Sheppard, Gurthwin Bosman, and R Heintzmann. Calculating point spread functions: Methods, pitfalls and solutions. *arXiv preprint arXiv:2301.13515*, 2023.

[4] Sarah Frisken Gibson and Frederick Lanni. Experimental test of an analytical model of aberration in an oil-immersion objective lens used in three-dimensional light microscopy. *JOSA A*, 9(1):154–166, 1992.

[5] Kyoji Matsushima and Tomoyoshi Shimobaba. Band-limited angular spectrum method for numerical simulation of free-space propagation in far and near fields. *Optics express*, 17(22):19662–19673, 2009.

[6] Marcel Leutenegger, Ramachandra Rao, Rainer A Leitgeb, and Theo Lasser. Fast focus field calculations. *Optics express*, 14(23):11277–11291, 2006.

[7] Reza Gholami Mahmoodabadi, Richard W Taylor, Martin Kaller, Susann Spindler, Mahdi Mazaheri, Kiarash Kasaian, and Vahid Sandoghdar. Point spread function in interferometric scattering microscopy (iscat). part i: aberrations in defocusing and axial localization. *Optics Express*, 28 (18):25969–25988, 2020.

[8] Craig F Bohren and Donald R Huffman. *Absorption and scattering of light by small particles*. John Wiley & Sons, 2008.

[9] Ernst Abbe. Beiträge zur theorie des mikroskops und der mikroskopischen wahrnehmung. *Archiv für mikroskopische Anatomie*, 9(1):413–468, 1873.

[10] Bahaa EA Saleh and Malvin Carl Teich. *Fundamentals of photonics*. john Wiley & sons, 2019.

[11] M Andreas Lieb, James M Zavislan, and Lukas Novotny. Single-molecule orientations determined by direct emission pattern imaging. *JOSA B*, 21(6):1210–1215, 2004.

[12] Simin Feng and Herbert G Winful. Physical origin of the gouy phase shift. *Optics letters*, 26(8):485–487, 2001.

[13] Gavin Young, Nikolas Hundt, Daniel Cole, Adam Fineberg, Joanna Andrecka, Andrew Tyler, Anna Olerinyova, Ayla Ansari, Erik G Marklund, Miranda P Collier, et al. Quantitative mass imaging of single biological macromolecules. *Science*, 360(6387):423–427, 2018.

[14] Daniel Cole, Gavin Young, Alexander Weigel, Aleksandar Sebesta, and Philipp Kukura. Label-free single-molecule imaging with numerical-aperture-shaped interferometric scattering microscopy. *ACS photonics*, 4(2):211–216, 2017.

[15] Shupei Lin, Yong He, Delong Feng, Marek Piliarik, and Xue-Wen Chen. Optical fingerprint of flat substrate surface and marker-free lateral displacement detection with angstrom-level precision. *Physical Review Letters*, 129(21):213201, 2022.

[16] Max Hantke and Gavin Young. Interferometric scattering microscopy methods and systems, October 27 2020. US Patent 10,816,784.

[17] Larnii S Booth, Eloise V Browne, Nicolas P Mauranyapin, Lars S Madsen, Shelley Barfoot, Alan Mark, and Warwick P Bowen. Modelling of the dynamic polarizability of macromolecules for single-molecule optical biosensing. *Scientific reports*, 12(1):1–15, 2022.

[18] B Th Thole. Molecular polarizabilities calculated with a modified dipole interaction. *Chemical Physics*, 59(3):341–350, 1981.

[19] Robert L Lucke. Fourier-space properties of photon-limited noise in focal plane array data, calculated with the discrete fourier transform. *JOSA A*, 18(4):777–790, 2001.

[20] Maurice George Kendall et al. The advanced theory of statistics. *The advanced theory of statistics.*, (2nd Ed), 1946.

[21] Bo Zhang, Josiane Zerubia, and Jean-Christophe Olivo-Marin. Gaussian approximations of fluorescence microscope point-spread function models. *Applied optics*, 46(10):1819–1829, 2007.

[22] Eugene Hecht. Optics 4th edition. *Optics 4th edition by Eugene Hecht Reading*, 2001.

[23] Jonathan Dong, Dante Maestre, Clara Conrad-Billroth, and Thomas Juffmann. Fundamental bounds on the precision of iscat, cobri and dark-field microscopy for 3d localization and mass photometry. *Journal of Physics D: Applied Physics*, 54(39):394002, 2021.

[24] John Jumper, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvunakool, Russ Bates, Augustin Žídek, Anna Potapenko, et al. Highly accurate protein structure prediction with alphafold. *Nature*, 596(7873):583–589, 2021.

[25] Yiwen Li, Weston B Struwe, and Philipp Kukura. Single molecule mass photometry of nucleic acids. *Nucleic acids research*, 48(17):e97–e97, 2020.

[26] Shuxiang Li, Wilma K Olson, and Xiang-Jun Lu. Web 3DNA 2.0 for the analysis, visualization, and modeling of 3D nucleic acid structures. *Nucleic Acids Research*, 47(W1):W26–W34, 05 2019. ISSN 0305-1048. doi: 10.1093/nar/gkz394.

[27] Claude E Shannon. A mathematical theory of communication. *The Bell system technical journal*, 27(3):379–423, 1948.

[28] Mahyar Dahmardeh, Houman Mirzaalian Dastjerdi, Hisham Mazal, Harald Köstler, and Vahid Sandoghdar. Self-supervised machine learning pushes the sensitivity limit in label-free detection of single proteins below 10 kda. *Nature Methods*, pages 1–6, 2023.