Best practice mass photometry: A guide to optimal single-molecule mass measurement

Jiří Kratochvíl[1,2], Raman van Wee[1,2,3], Dan Loewenthal[1,2], Jan Christoph Thiele[1,2], Jack Bardzil[1,2], Kishwar Iqbal[1,2], Stephen Thorpe[1,2,*], Philipp Kukura[1,2,*]

1) The Kavli Institute for Nanoscience Discovery, University of Oxford, Dorothy Crowfoot Hodgkin Building, South Parks Road, Oxford OX1 3QU, UK

2) Physical and Theoretical Chemistry Laboratory, Department of Chemistry, University of Oxford, South Parks Road, Oxford OX1 3QZ, UK

3) Department of Physiology, Anatomy and Genetics, University of Oxford, South Parks Road, Oxford, OX1 3QX, UK

*To whom correspondence should be addressed: stephen.thorpe@chem.ox.ac.uk, philipp.kukura@chem.ox.ac.uk

Key points

The protocol optimizes the interplay between surface and sample preparation, experimental technique and, analysis parameters to maximum sensitivity and resolution when performing MP instruments.

All of the changes can be implemented into MP workflows with less than a day's work and many are only simple considerations at the time of the measurement itself.

Key references [Up to 5 articles relevant to this protocol; where the method has been introduced and/or used.]

R.Asor et al. *Proc. Natl. Acad. Sci. U.S.A*. 121 (40) e2403260121 (2024). https://doi.org/10.1073/pnas.2403260121
A.Sonn-Segev et al. *Nat Commun* 11, 1772 (2020). https://doi.org/10.1038/s41467-020-15642-w
A. Olerinyova et al. *Chem*, Volume 7, Issue 1, 224 - 236
G. Young et al. *Science* 360,423-427(2018). https://doi.org/10.1126/science.aar5839

Abstract

Mass photometry (MP) has emerged as a powerful approach to study quaternary biomolecular structure, dynamics and interactions in as little as a few minutes under native conditions. The capabilities of the method ultimately hinge on the ability to accurately measure the tiny optical contrast generated by individual molecules at a glass-water interface, which enables mass-resolved quantification of biomolecular mixtures. Ideally, this capability is only limited by shot noise, but in practice depends on additional parameters and details of the assay. Here, we focus on the key factors affecting MP performance and present simple steps that can be taken to achieve optimal MP performance in terms of mass resolution, quantitative detection limit, reproducibility, and analyte concentration range without compromising the speed and simplicity of the technique.

[H1] Introduction

Mass photometry (MP) is a label-free optical method for the solution-based mass measurement of single biomolecules[1,2]. The measurement principle relies on quantifying the change in surface reflectivity as a result of the presence of individual biomolecules and their complexes, which is proportional to the molecular mass. MP has been used to study a variety of phenomena, such as protein oligomerisation[3–10], adeno-associated virus assembly[11,12], protein-DNA[13–16], protein-protein[17], protein-ligand[18,19] and antibody-antigen interactions[20,21], among many others[22–24]. In addition to the original implementation relying on non-specific adsorption of the analyte to a bare glass surface, dynamic measurements of analyte movement on the surface of lipid-bilayers[25,26], or measurements of proteins at high concentrations using polyethylene glycol (PEG) passivation surfaces are also possible[27]. Previous protocols for MP have focused on the basics of performing an MP measurement with a focus on binding affinities[28,29].

Standard MP experiments require the addition of as little as 4-20 µl of sample to a microscope coverslip, and analysis can take as little as a few minutes. The resulting data provide immediate information on sample composition, resolved by molecular mass. Owing to its simplicity, MP is extremely robust and broadly applicable. At the same time, minimal changes to the assay can dramatically improve the final data quality. By systematically exploring the influence of noise sources affecting single molecule measurement precision, we illustrate and define a workflow that mitigates potential pitfalls and thereby enables easy optimization of MP measurements. As a result, we enable the user to routinely obtain the best possible data quality without affecting the simplicity and speed of MP measurements. For analytes that do

not bind well to glass, such as nucleic acids[29] we show that (3-aminopropyl)triethoxysilane (APTES) functionalization to create positively charged glass through surface amination[30–33] can improve binding and thereby increase the measurement precision and concentration range.

[H2] Principle of MP

The basic detection principle of MP rests on quantifying the difference in the reflectivity of an interface in the presence and absence of a biomolecule. Most commonly, that change originates from the interference between light scattered by the biomolecule binding to a glass-water interface, and light reflected by the same interface (**Fig. 1a**). The use of a thin metal film amplitude mask placed in the back focal plane selectively attenuates reflected light without affecting the scattered light, enhancing the optical contrast and enabling much higher power density at the sample without saturating the detector, thereby improving the experimentally achievable signal-to-noise ratio (SNR) for a given exposure time[34].

The raw image of the glass-water interface is dominated by scattering caused by the nanoscopic roughness of the glass substrate, which results in a speckle-like pattern (**Fig. 1b**)[35]. Depending on the illumination wavelength and mask strength used, the resulting peak-to-peak intensity variation is on the order of 10-30% of the reflected light intensity. Binding of a 180 kDa protein dimer of dynamin 1 with its proline-rich domain deleted (Dyn1 ΔPRD) to this surface results in a small change in the reflectivity, in this case on the order of 0.5%. Such a change in reflectivity is challenging to detect using digital cameras, because of fluctuations in the background light intensity resulting from the Poissonian behaviour of photon arrival on the detector (shot noise). The effects of shot noise are, however, substantially reduced if temporal averaging is applied, making it possible to not only detect, but also quantify the intensity change (**Fig. 1c**).

While this step change is easily visible when zoomed in on a single camera pixel, it is undetectable in the raw image, because the signal magnitude (0.5%) is 1-2 orders of magnitude smaller than the image speckle caused by the glass surface roughness.

The lack of visibility can be addressed by computing ratiometric images as a function of time. This is done by calculating the ratio between two consecutive windows where each window is made up of multiple frames. Averaging over multiple frames reduces the impact of shot noise-induced background fluctuations (**Fig. 1c**) while the division removes the stationary image speckle caused by glass surface roughness, enhancing visibility and simplifying detection (**Fig. 1d**). For each pixel, the normalized contrast is given by Eq. 1

$$C_i = \frac{\sum_{j=1}^{n} x_{i+j}}{\sum_{j=1}^{n} x_{i-j}} - 1 \quad \text{(Equation 1)}$$

where *n* is the window size in frames and *x* the raw frames recorded by the camera. Computing the ratiometric image for a binding event at the maximum of this trace yields an image of the binding event, whose visibility is strongly affected by the degree of spatial and temporal binning, which in turn affects shot noise-induced background noise (**Fig. 1e**). The total averaging time for an MP image is given by the camera exposure time multiplied by the number of temporally binned frames.

In an experiment where the analyte is present in solution, proteins continuously bind to the glass surface. The contrast for each binding is quantitively measured by fitting the image of the binding event to a point spread function model, the amplitude of which gives the binding event contrast. Given sufficient single molecule contrast measurement precision, the contrast variation from event to event is small enough, such that species of different mass can be

resolved, resulting in clear stripes when plotting the detected contrast as a function time (**Fig. 1f**). Integrating these events into a contrast histogram yields baseline-resolved peaks at integer multiples of the monomer contrast (**Fig. 1g**). To convert from contrast to mass, a known oligomeric calibrant protein or mixture of monodisperse proteins is measured initially. The linear dependence on contrast with protein volume means that essentially any protein can be used as a calibrant providing the expected molecular masses can be resolved. Plotting the center of mass of each peak *vs* the known molecular mass of the different calibrant oligomers yields an accurate linear mass-to-contrast relationship (**Fig. 1h**), where we allow for a y-offset to improve the fit[1]. This calibration can then be applied to unknown samples, which together with the high achievable mass resolution, enables identification and quantification of biomolecular mixtures by mass (**Fig. 1i**).

The key performance metrics for MP measurements are:

(1) Mass accuracy; the difference between the measured and expected mass;

(2) Mass resolution; the smallest mass difference for which two species in a mixture can be resolved;

(3) Quantification; to which degree the measured molecular counts are proportional to the abundance of each species in the mixture;

(4) The lower mass detection limit. For optimal measurements, these values have reached 2% accuracy, 20 kDa resolution and a quantitative detection limit on the order of 40 kDa.[1]

Here, we identify and quantify noise sources and experimental considerations, show how they affect these key metrics, and provide guidelines for how to achieve optimal data quality in routine measurements.

[H2] Repeatability and dynamic range

To characterise the reproducibility of MP measurements, we consecutively measured eight replicates of *Dyn1 ΔPRD* oligomers at a monomer-equivalent concentration of 80 nM (note that for predominantly monomeric proteins the optimal sample concentration will be lower), (**Fig. 2a)**[25]. We chose this molecule because of its high mass purity and propensity to oligomerise. To achieve maximum reproducibility in the total counts, it is important to ensure that the incubation time in the sample tube and the time between sample addition and the beginning of the measurement is consistent, as both the sample tube and the measurement coverslip will exhibit time dependent protein adsorption. Under such conditions, the total event count across eight replicates fluctuates by about ±15% around the mean of the total event counts (Figure 2a). This number can differ depending on how and where the analyte is added to the measurement gasket and convection within the droplet. The monomer:dimer ratio, exhibits a smaller fluctuation on the order of ± 5% (**Fig. 2b**). This ratio is less sensitive to how the sample is added and instead depends on sample concentration and incubation time through the relevant interaction affinities.

The retrieved mass-to-contrast slope from a linear fit as shown in **Fig. 1h** can vary by as a little as ±1% when care is taken in terms of sample addition (**Fig. 2b, see procedure**). This observation is confirmed when examining the ratiometric contrast of the different oligomers (**Fig. 2c**); in each case the normalised contrast varies by less than 2% when oligomers of different mass are compared. This variation limits the smallest changes detectable by MP, where complete analyte population shifts from mass m to mass m+2%, which can be as low as 1 kDa[1].

We emphasise the difference here between measurement precision, which is sufficient for visualising binding of molecules below the kDa level[1], and resolution, which requires the simultaneous and independent observation of two species. In general, the repeatability in terms of mass measurement is most sensitive to the sample focus, which can be reliably optimized by maximising the optical contrast of the glass roughness in the raw image.

An important aspect of MP that has been largely neglected to date is the dynamic range in terms of relative abundance of species present in the mixture. Given that single biomolecules or their complexes are detected in a digital fashion, the dynamic range is in principle only limited by measurement time and field of view (FoV) size. For the example shown here, a linear representation of the data reveals monomer, dimer, tetramer, hexamer and small amounts of octamer. Plotting the same data on a log scale (**Fig. 2d**) shows clear signatures of decamers and dodecamers, as well as low counts of even larger species. In practice, the dynamic range is often limited by impurities in the sample, but if these can be avoided and total counts exceed 10^4, it can reach $>>10^3$, as shown by the reproducible detection of a species with only 10 counts.

[H2] The impact of detection threshold, averaging time, and field of view

In the absence of other noise sources, the ability to detect and quantify individual biomolecules by MP is only limited by shot noise-induced fluctuations of the imaging background from reflected light at the glass-water interface. For shot noise, the expected fluctuation for the detection of N photoelectrons per pixel is $\sqrt{N}$, thus the magnitude of background fluctuations is given by $\sqrt{N}/N = 1/\sqrt{N}$. Detecting more photoelectrons per unit

time thus reduces the image noise and improves detection and quantification of individual biomolecular events.

The incident power is limited first by the damage threshold of the optical components and sample, and second by the finite full well depth and frame rate of the detector. Increasing the temporal averaging window means more frames can be acquired, thereby increasing N, reducing shot noise and improving the SNR of the PSF (**Fig. 3a**), which improves the detectability of particles (**Fig. 3b**). The same effect is achieved by over-magnification of the image beyond the Nyquist limit across the detector and subsequent spatial binning to combine the full well depths of multiple pixels, allowing higher photon counts without detector saturation.

Quantitative detection is critical for accurate characterisation of affinity between biomolecules. An experimental approach to evaluate quantitative detection involves titrating towards zero analyte concentration: if no false positives are detected, a concentration vs counts plot should exhibit a linear dependence with an intercept near zero. For a high SNR protein such as a 150 kDa anti-PSMC6 antibody (Abcam RRID:AB_447214) at 100 ms averaging time, such behavior is indeed observed experimentally (**Fig. 3c**). Repeating the experiment for the same averaging time and analysis parameters for protein A (42 kDa), reveals a slight positive offset for the intercept, indicative of a small amount of noise features being counted for the given detection parameters (**Fig. 3c**). If required, one can iterate the acquisition and analysis parameters to optimise the performance as a function of analyte mass.

Signal strengths that approach the noise limit also influence mass accuracy in the low mass range. A low detection threshold will lead to substantial false positives, while a high detection threshold results in a loss of low SNR events. This will lead to an effective hard cut off on the

low mass end of the distribution, thereby shifting the measured mass to higher values (**Fig. 3d**). This effect can be observed when characterising proteins in the 27 to 66 kDa mass range. As previously, this effect can be minimized by simultaneous optimization of averaging time and analysis thresholds, where required. In essence, this scenario will lead to an underestimate of the number of species and an overestimate of their mass, because only the high mass side of the distribution is detected. As a result, the mass-to-optical contrast linear relationship will drop off too slowly as lower mass species are detected.

Mass resolution between species of similar mass in a sample assay is influenced by many factors. The first factor is a high enough SNR to accurately estimate the individual peak contrast, with the full width at half maximum (FWHM) reducing from 31 kDa to 19 kDa with increased frame averaging for the tetramer peak of Dyn1 ΔPRD shown here (**Fig. 3e**). However, the absolute mass resolution is generally lower for larger species because relative deviations in event contrast due to locally varying differences in glass roughness as well as the larger absolute error of minor discrepancies in the PSF fit. This results in an increased fitted peak standard deviation (sigma, σ) for higher mass species. (**Fig. 3f**).

The dependence of mass resolution on SNR can also be observed by its dependence on the size of the FoV, reducing *N* as optical power density is distributed over a larger area (**Fig. 3g**). Loss, rather than continuing gain of mass resolution for longer averaging times arises from additional, non-shot noise sources.

[H2] Non-shot noise contributions

The results in **Fig. 3** demonstrate that several key performance parameters of MP can be improved by increasing the averaging time, which in turn reduces shot noise-induced background fluctuations. However, if this relationship were to hold indefinitely, there should

be no limits to detection sensitivity or mass resolution for MP. We can test this relationship by computing signal fluctuations in the absence of analyte in both space and time. To do this, we evaluate the contrast standard deviation of a series of images on a pixel-by-pixel basis (**Fig. 4a**), which can be converted to a mass-equivalent using a mass-to-contrast calibration (**Fig. 1h**). For a purely shot noise-limited case, we would expect this noise to scale with the inverse square root of the averaging time.

With increasing averaging time, the single image (**Fig. 4b, top FoV**) pixel noise drops, following the expected 1/√N dependence, and reaches a minimum of 7.1 kDa at 128 ms (**Fig. 4b, middle FoV**). For longer averaging times, a dynamic speckle-like background appears, and the pixel noise increases, deviating from the expected shot-noise limit (**Fig. 4b, bottom FoV**). The source of this excess noise is currently unknown, but we have previously ruled out lateral sample drift (on sub-second timescales) and shown that excess noise cannot be removed by temporal averaging[36]. We have observed a similar noise limit across many different instruments, buffers, and averaging times, suggesting that its source is connected to the measurement itself, rather than being the result of an imperfect experimental approach. In practice, we have found that choosing an averaging time about 20% shorter than the noise floor determined by this plot produces the overall best compromise between detection limit and mass resolution. Optimizing the averaging time along with the correct threshold of minimum contrast change for a particle to be detected is critical to minimize false positives whilst still detecting low molecular weight species.

In contrast to protein binding events, the spatiotemporal nature of the non-shot noise (**Fig. 4c**) results in its detection as an equal number of events on the positive and negative mass axis at the same mass. The number of events within these symmetrical peaks increases for a given

averaging time as the detection threshold is lowered. The appropriate threshold can be determined empirically by adjusting the analysis parameters to obtain <100 negative mass events in total when making a recording of pure measurement buffer prior to protein addition.

[H3] **Common causes for deviation from optimal performance**

Having characterized the fundamental noise contributions associated with MP, we turn to the most common causes for deviations from optimal performance. A relatively trivial, but often encountered effect arises from sub-optimally cleaned microscope glass coverslips. Remaining impurities evidence themselves as high-contrast particles on the surface in the raw image (**Fig. 4d**). The signal generated by these large particles can easily saturate the camera leading to inaccurate mass quantification of any biomolecules landing on top of them. Pixel noise is also clearly elevated relative to a clean coverslip (**Fig. 4e**), likely due to the imperfect ratiometric removal of the large and fluctuating associated signals generated by motion that spans multiple pixels. Hence, we recommend moving to a different FoV without these large particles as evidenced by reduced image sharpness or repeating the coverslip cleaning procedure if a brief search for a clean area is unsuccessful.

Increased background fluctuations can also be generated by dissolved macromolecules whose mass is below the detection limit, such as poly-ethylene glycol (PEG) used for molecular crowding if present in high abundance in the buffer (**Fig. 4e**)[37]. The background fluctuations generated by these macromolecules results in correspondingly lower SNR and increased false positive detection.

In addition, back reflections from the water-air interface of the droplet can produce large fringes in the ratiometric images (**Fig. 4f**), which can be mitigated by adjusting buffer volume and the lateral position of the sample.

Finally, depending on their purity, biological samples can contain large particles whose landing on the coverslip generates a signal well above the expected background signal (**Fig. 4g**). These can easily be excluded from the analysis as required by masking the affected spatiotemporal components of the recording.

On top of these effects that can affect MP measurements generally, there are a few specific ones that contribute to a loss of mass resolution. An easily identifiable one is focus drift, which can occur when adding chilled analyte to a coverslip, which is at room temperature. A change in temperature leads to differences in local refractive index, affecting the focus and optical contrast. Such behavior can be easily spotted and removed by plotting the contrast as a function of measurement time, mitigating unwanted mass broadening. In a standard landing assay, focus drift values are around 0.5% of the peak mass per minute (**Fig. 4h**). In some cases, non-uniform illumination and glass roughness can also cause local reflectivity variations which manifest themselves as differences in a local mass-to-contrast conversion. As a result, events in different parts of the FoV exhibit systematically different contrasts. This effect can be mitigated by only analysing parts of the FoV, comparing mass histograms from different sub-FoVs, or using reflectivity correction (**Fig. 4i, j**).

[H2] Characterization of the quality of binding events

The fundamental concept of MP relies on accurately quantifying the change in reflectivity associated with a biomolecule binding to the glass-water interface. **Figs**. **3** and **4** illustrated the potentially detrimental influence of excessive background fluctuations caused by either shot noise or other sources. In addition, the dynamics of individual landing events themselves can cause an error in the extracted particle contrast.

The most commonly encountered protein dynamics near the interface observed by MP are illustrated in **Fig. 5a**.

- *Binding events* – the protein lands and remains stationary on the surface, indicating stable and optimal binding.
- *Unbinding events* – the protein temporarily adheres to the surface and then detaches.
- *Rolling events* – proteins that land on the surface followed by lateral movement, resulting in a distinct binary event with a binding head and unbinding tail in the ratiometric image.
- *Wobbling events* – proteins sometimes adsorb weakly to the surface and proceed to exhibit erratic, multidirectional movement.

The latter two event types can sometimes be erroneously identified by the particle picking algorithm as multiple events.

Any of the characteristics that deviate from a regular binding event can be considered as a suboptimal event. This arises from the need to average frames to reduce shot noise: if the particle departs or moves during the averaging time, it will decrease the local reflectivity change, causing a contrast and thus mass measurement error. The principle behind this effect is most easily understood when considering the consequences of rapid unbinding on the intensity of an individual camera pixel (**Fig. 5b**). Ratiometric processing of the single pixel trace of BSA dimers reveals a typical contrast of 0.35% when averaged over all optimally binding events (**Fig. 5c**). In the example with subsequent unbinding, however, the event yields a contrast of 0.23%, over a third less than expected. This contrast reduction is caused by the inclusion of the period without the particle present on the surface during the 80 ms averaging

time. Although the impact of other suboptimal event types may be more subtle, one can easily understand how they lead to similar effects on the single (sub-diffraction) pixel level.

To mitigate the effects of suboptimal binding and false positive detection, one can use simple filtering approaches in post processing. For example, noisy frames with increased background noise evidenced by spikes in standard deviation, such as those caused by large aggregates, can be excluded from analysis (**Fig. 5d, red and Fig. 5e**). The same frame (**Fig. 5e**) also shows signatures from wobbling particles, repeated detection of which results in regions with high landing density, that can be excluded from analysis as well (**Fig. 5f, red**). The substantial presence and cumulative impact of poor binding events can distort the mass histogram, as for an aged stock solution of Dyn1 ΔPRD (**Fig. 5g, red**). This distortion manifests itself in several ways: the emergence of an unbinding peak, i.e., a mirrored negative mass peak; the broadening of peaks, frequently accompanied by a characteristic shoulder on the low mass end; and an increase in the mass histogram baseline noise between peaks. These distortions complicate the interpretation of MP data, especially when coupled with the aforementioned noise sources, underlining the importance of accurately characterising and mitigating suboptimal binding events. By employing spatial and temporal filtering shown in **Fig. 5d–f**, a corrected mass distribution can be recovered (**Fig. 5g, black**), achieving performance comparable to that of an optimally binding sample (e.g., **Fig. 1i**). This underscores the critical role of mitigating poor sample binding in optimizing MP performance.

[H2] Improving binding quality

An ideal measurement consists only of binding events. However, depending on the charge of the glass surface and the biomolecule of interest, binding can be reversible. In scenarios where

the affinity of the biomolecule to the surface is low, binding and unbinding events are observed even at low levels of surface coverage. This can lead to deterioration of measurement performance in terms of both mass accuracy and resolution (**Fig. 6a,b**). In general, for a high binding:unbinding ratio a linear scaling of total counts with concentration and optimal mass resolution is achieved (**Fig. 6c**). For higher concentrations, excessive unbinding breaks the linear scaling of counts as unbinding particles increase the local concentration above the surface, and thereby the overall binding rate.

For BSA, functionalising the coverslip surface with positively charged APTES increases the surface affinity of the protein (BSA is negatively charged at neutral pH), evidenced by the near doubling of molecular counts for the same analyte concentration and a drop in unbinding counts (**Fig. 6a,d**). The improvement in protein binding to the surface markedly increases the mass resolution at a concentration of 20 nM (**Fig. 6b**,**e**). The corresponding titration exhibits similar resolution and slightly improved count scaling with concentration in the 1-10 nM analyte concentration range (**Fig. 6f**). In the 10-50 nM concentration range, however, APTES functionalization improves those parameters compared to those achievable on a bare glass surface, because of reduced unbinding that would otherwise contribute to overcrowding of the FoV (**Fig. 6c**,**f**). APTES functionalization thus elevates the upper concentration limit in this case about five-fold with otherwise identical performance. This is important, for example, for the quantification of protein-protein interactions with lower affinity as it enables measurement at concentrations closer to the dissociation constant[27]. At 100 nM, the count linearity in concentration breaks down even on APTES, due to overcrowding of the FoV and saturation of the surface. The effect of protein surface saturation can also manifest itself during the measurement, as more proteins bind to the surface. For example, at 50 nM, low levels of unbinding are observed during the first minute of the measurement (**Fig. 6g**). After 3

minutes however, the APTES surface becomes saturated, and a substantial unbinding peak emerges (**Fig. 6h**). The temporal evolution of binding and unbinding demonstrates that optimal binding conditions are only satisfied for the first minute of the measurement, beyond which the measurement quality degrades (**Fig. 6i**). These results illustrate the advantage of starting the measurement immediately after sample addition to avoid unnecessary surface coverage by the analyte.

[H2] Importance of binding quality

Previous sections discussed the effect of binding properties on measurement performance and resulting histograms. In summary, unbinding and poor binding quality can pose issues with:

- Accurate mass estimation: Binding and subsequent unbinding within the averaging window leads to mass underestimation and peak broadening (**Fig 5c**).
- Low mass measurement: The unbinding peaks (**Fig. 6b**) can interfere with the noise peak (**Fig. 4c**), which makes distinguishing signal from noise difficult.
- Quantification of interactions: Species with different unbinding rates influence binding peaks ratios. A typical example is a mixture of a well-binding protein and an antibody with excessive unbinding, which results in higher local antibody concentration above the surface and an exaggerated antibody binding peak.
- Interactions with lower affinity: Unbinding causes premature FoV overcrowding and prevents quantification at higher concentrations.

Poor binding events can be minimized by lowering the measurement concentration, surface functionalization, or reducing the measurement duration (**Fig. 6i**). If poor binding events

cannot be avoided, they can be filtered out in the analysis (**Fig. 5d-g**). We emphasize that with consistent sample addition and a good binding surface, total counts can indeed be quantitative and scale with analyte concentration, with a slope fit error of only 1.1% (**Fig. 6f**).

[H2] Quantifying species abundance

Whether an analyte can be detected by MP depends on both its mass and the averaging time, because they together define the signal-to-noise-ratio. As event contrast scales linearly with mass up to 5 MDa (potentially 10 MDa for particles with compact shapes), it is intuitive that the SNR decreases for smaller species for a constant averaging time (**Fig. 3b**).

Increasing the averaging time improves detection at low mass, which directly translates into a more accurate monomer:dimer ratio as shown here for BSA (**Fig. 7a**). Increasing the averaging time further may improve mass resolution but has no impact on the total number of detected events or the monomer:dimer ratio. This suggests that the peak ratio is quantitative for averaging times >20 ms.

At 10-fold higher analyte concentration however, longer averaging times lead to a drop in total detected events and the monomer:dimer ratio (**Fig. 7b**). In addition, the monomer:dimer ratio drops from 17 to 12 as expected for a higher measurement concentration. Event crowding makes it more challenging to detect low contrast events whose PSF partially overlaps with that of larger species, an effect that becomes more prominent with increasing averaging time (**Fig. 7c**). This, together with the need for sufficient averaging time to detect low mass species leads to a clear maximum of both the monomer:dimer ratio and the number of detected events and thus defines the optimum averaging window of 20 ms (**Fig. 7b**).

[H2] Importance of concentration for measurement quality

The previous sections have shown the influence of protein concentration and the resulting landing density that appears in ratiometric images. To summarise, too high concentrations can cause issues with:

- Protein binding: Saturated surfaces reduce binding quality, resulting in unbinding, wobbling, and rolling molecules.
- Quantitative detection: Overlapping PSFs will not be reliably detected, reducing molecular counting accuracy.
- Contrast accuracy: Overlapping signals cause errors in contrast and thus mass measurement.
- Mass sensitivity: Overlapping PSFs are exacerbated at high averaging times. Increasing averaging to detect lower mass species will be ineffective.

As the particle landing density is also a function of protein oligomerisation and affinity to the surface, one cannot define a single concentration that can be used for every protein. To obtain optimal MP performance, a titration to find the optimal concentration range is helpful for each analyte.

[H2] Overview of the procedure

To mitigate potential challenges that can be encountered when making an MP measurement and optimize performance we provide a step-by-step protocol for preparation, measurement and analysis (**Fig. 8**).

To demonstrate the efficacy of the protocol we detail experimental conditions for two easily accessible proteins, BSA and SARS-CoV-2 antibodies. BSA serves as an example of a poor

binding low mass oligomeric protein, which is challenging to measure with high accuracy and reproducibility, while the SARS-CoV-2 antibody is included, because of the community's interest in using MP as a technique to characterise antibody-antigen interactions.

For APTES functionalization, we selected a variation of the protocol that we developed for surface passivation for MP by PEGylation[26] as a simple approach for NH_2-rich coatings without aggregates. Any other APTES protocols can be used[29-32] if they provide an optically clean and sufficiently charged surface. Nonetheless, this protocol is intended as a starting point for measuring a protein of interest and the described optimization strategy can be directly applied to other samples.

The protocol has been written based on currently available commercial instruments as these are the most used by the MP community. All discussed principles and much of the protocol are directly applicable to custom instruments. The absolute performance in terms of sensitivity and resolution will be determined by the instrument itself but the optimum data quality can still be achieved using this same approach. The only changes required will be in the data analysis if the user is relying on custom software.

[H3] Expertise needed

To perform the MP measurements of the protocol, basic knowledge of how to use a mass photometer using the AcquireMP software and how to perform analysis using the DiscoverMP software is required. Anyone with experience with standard wet lab techniques will be able to perform the remainder of the protocol without difficulty.

[H2] Limitations

1. **Detergents**: Detergent micelles in themselves are protein-sized objects and thus produce comparable signal that creates substantial background. This can be mitigated using special detergents, or peptidiscs.[4,19]
2. **Protein mass:** The protein mass and the mass differences between biomolecules complexes need to be larger than 30 kDa for quantitative measurements.
3. **Buffer and protein purity**: Proteins and buffer will ideally be free of any other particles. High concentrations of low mass particles, such as PEG or glycerol below the detection limit can cause additional noise (**Fig. 4e**). Note that buffer ions are sufficiently small to not contribute to background noise as long as buffers are prepared well in the solubility region of salts. Low abundance contaminants above the detection limit can be removed from the mass histogram by spectral decomposition, provided they do not participate in the interaction(s)[33].
4. **Low-affinity interactions**: Low affinity complexes or oligomers with are difficult to quantify as they dissociate at the low measurement concentrations required for MP. In the case of slow off-rates, incubation at high concentration (>1 µM) and rapid manual dilution can overcome the issue, provided the off rate is sufficiently slow for the complex(es) not to fall apart[21,38]. However, complexes with a fast off-rate on the

order of seconds or less, and a dissociation constant above 100 nM are difficult to measure using standard landing assays. This can be overcome by using robustly passivated surfaces with nanoholes accessible for binding to reduce the landing rate, which enables measurements at higher concentrations in equilibrium, or by exploiting rapid dilution using microfluidics[27,39].

[H1] Materials

[H2] Reagents

- Gibco™ DPBS, no calcium, no magnesium (Fisher Scientific, cat.no. 15326239)
- 3-Aminopropyltriethoxysilane 99% corc-sealed (Sigma-Aldrich, cat.no. 440140)
- Acetone 99.8%, HPLC grade (Sigma-Aldrich, cat.no. 270725)
- 2-propanol ACS reagent, ≥99.5% (Sigma-Aldrich, cat.no. 190764)
- Bovine serum album lyophilized powder (Sigma-Aldrich, cat.no. B6917)
- SARS-CoV-2 (2019-nCoV) Spike Neutralizing Antibody, Mouse Mab (SinoBiological, cat.no 40591-MM48 RRID:AB_3677184)
- Massference-p1 (Refeyn MP-CON-41033)
- MilliQ water, >18 MΩcm, 0.22-µm filtered, ultrapure (type 1) water
- Nitrogen N5.5 99.9995% Research Grad (Boc)

[H2] Equipment

- TwoMP (Refeyn)
- AcquireMP Version 2.5 (Refeyn)
- DiscoverMP Version 2023 R1.2 (Refeyn)

For using Refeyn consumables:

- Sample Preparation Kit (*Optional*, MP-CON-21008, Refeyn)
- Samples well cassettes (6 wells) for the OneMP, TwoMP, SamuxMP (MP-CON-41005, Refeyn)
- Nanodrop (Thermo Scientific, Nanodrop 1000 Spectrophotometer)

For cleaning own coverslips:

- Carbon PEEK Replaceable Tip Tweezers (Agar Scientific, cat.no. AGT551)
- Fisherbrand™ Borosilicate Glass Tall Form Beakers 400ml (Fisher Scientific, cat.no. 15439093)
- Epredia No 1.5 coverslips 50 x24mm (VWR, cat.no. 16002-264)
- Slidebox for 25 Slides (GLW, cat.no. K25W)
- Optical Grade Cotton-Tipped Applicators (Thorlabs, cat.no. CTA-10)
- Carl Zeiss Immersol 518 F Immersion Oil (Fisher Scientific, cat.no. 10539438)
- Zepto Plasma Cleaner (Diener Zepto-BRS 200)
- Epredia™ Coverglass Staining Rack (Fisher Scientific, cat.no. 12627706)
- Ultrasonic Cleaning PROCLEAN 10.0S (Ulsonix)
- Lens Cleaning Tissue (Thorlabs, cat.no. MC-5)
- Velp Scientifica™ AREC.X 7 Digital Ceramic Hot Plate Stirrer (Fisher Scientific, cat.no. 17274083)
- Magnetic stirrer bar, 8 mm x 3mm (Fisher Scientific, cat.no. 11818892)
- RS PRO Infrared Thermometer (RS, cat.no. 136-7890)
- B Braun™ Injekt Solo Cone Syringes (Fisher Scientific, cat.no. 12722637)
- B Braun™ Sterican 70mm 20-gauge needle (VWR, cat.no. 612-0159)
- B Braun™ Sterican 16mm 25-gauge needle (VWR, cat.no. 612-0153)

- Silica Gel (RS, cat.no. 388-8421)
- Eppendorf® PCR tubes (Sigma-Aldrich, cat.no. EP0030124537)
- Eppendorf® Flex-Tubes Microcentrifuge Tubes (Sigma-Aldrich, cat.no. EP02236411)
- SalvisLab Vacucenter Vacuum Oven (Cole Parmer, cat.no. WZ-52402-36)

[H2] Reagent setup

[H3] Massference-p1 preparation

Timing 30 minutes

1) Defrost a vial of Massference-p1.
2) Aliquot into 2 µl vials.
3) Freeze at -20°C and store until required.

[H3] BSA preparation

Timing 30 minutes

1) Prepare a 60 µM solution in Dulbeccos PBS. When dissolving lyophilized BSA powder in PBS vortex slowly to avoid generation of bubbles.
2) Verify concentration using Nanodrop.
3) Store at 4°C until measurement, can be stored for up to 1 week.

[H3] Antibody preparation

Timing 10 minutes

1) Prepare an 8.3 µM of SARS-CoV-2 antibody solution in Dulbeccos PBS.
2) Verify concentration using Nanodrop.
3) Store at 4°C until measurement, can be stored for up to 1 month.

[H2] Equipment Setup

[H3] Monthly instrument calibration (optional)

Timing 2.5 hours

1) Ensure the microscope objective is completely clean before starting. Clean with IPA, optical tissues and cotton buds if not.
2) Switch on the mass photometer and open the acquisition software (AcquireMP) 2 hours before measurement to ensure thermal equilibrium.
3) Place an array of six gaskets at the center of a cleaned coverslip.
4) Put a drop of oil onto the objective and place coverslip onto microscope stage. Ensure sufficient, but no excess oil.
5) Add 20 µL of MilliQ in the gasket, wait for 1 minute with a closed lid to equilibrate temperature.
6) Focus on the glass-water interface.
7) Go to 'Acquisition' in the DiscoverMP menu bar in the top of the screen and select 'calibrate the acquisition image' and 'calibrate run times', to ensure homogeneous illumination and minimize scanning noise.

[H1] Procedure

Stage I: Coverslip cleaning

Timing 30 minutes

<CRITICAL> If using precleaned coverslips from Refeyn, go to Stage III: Measurement preparation.

1) Place coverslips in a cleaning rack with tweezers or gloved hands, avoid touching the measurement area.
2) Clean coverslips by sonicating in acetone, 1:1 MilliQ:Isopropanol solution and MilliQ, each for 5 minutes.
3) Blow dry coverslips with nitrogen and store them in a slidebox.

TROUBLESHOOTING

Stage II: Surface amination (optional)

Timing 2 hours

<CRITICAL> If analyte binds well to glass, go to Stage III: Measurement preparation.

4) Pour 250 mL acetone into a beaker, cover it with aluminum foil or a glass Petri dish to avoid solvent evaporation and heat the solution to 45°C (measured by an optical thermometer). To mitigate safety risks, always handle acetone in the fume hood, perform heating using the water bath and avoid any sparks or open flames.
5) Add 5 mL of APTES to the beaker, using a syringe with an attached needle, to obtain a 2% solution.
6) To activate the glass surface with -OH groups for the subsequent silanization, insert cleaned coverslips into the oxygen plasma cleaner.
7) Pump the chamber to a pressure below 0.15 mbar.
8) Introduce O_2 to reach 0.8 mbar pressure.
9) Plasma-clean the coverslips for 8 minutes at 50% power.
10) After plasma cleaning, promptly put the coverslips into a beaker with acetone and then immediately transfer into the beaker with 2% APTES solution.
11) Incubate coverslips with APTES at 45°C for 15 min with magnetic stirring.

12) Sonicate APTES beaker with coverslips for 1 minute to remove physiosorbed APTES molecules and aggregates from the surface of beaker walls.
13) Incubate for another 15 minutes while stirring at 45°C. Sonicate the coverslips in fresh acetone twice for 5 minutes, replacing the acetone between each sonication and then sonicate for 1 minute in MilliQ.
14) Dry APTES coated coverslips with nitrogen and store in a slidebox with silica gel.
15) Clean APTES beakers by removing the 2% APTES solution and sonicating in fresh acetone for 30 minutes.

Critical Step. Using low-water content acetone is important to avoid APTES polymerization and aggregate formation, which increases background noise. A well-prepared APTES surface should be free of aggregates with a comparable raw image to glass (**Fig 1b**).

Critical step. Quickly transferring the coverslips from the plasma cleaner to the APTES beaker is essential as -OH group activation decreases with time.

Pause point: For best performance, use the coverslips on the day of preparation but they can be stored under cover at room temperature (20-25 °C) for up to 1 week.

TROUBLESHOOTING

Stage III: Measurement preparation

Timing 1 hour

16) Ensure microscope objective is completely clean before starting. Clean with IPA, optical tissues and cotton buds if not.
17) Switch on the mass photometer and open the acquisition software (AcquireMP) 1 hour before measurement to ensure thermal equilibrium.

18) Prepare an Eppendorf with a 1.5 mL aliquot of buffer and equilibrate buffer at room temperature.

19) Place six gaskets onto the center of a plain or functionalized coverslip.

20) Put a droplet of oil on the objective and place the coverslip on the microscope stage. Ensure sufficient but not excess oil.

21) Select FoV according to intended analyte.

- Small for <70 kDa (e.g., BSA).
- Regular for ≥70 kDa (e.g., SARS-CoV-2 antibody).
- Large for measurement of large species with low abundance.

Critical step. To enable data processing with short averaging times by setting AcquireMP to advanced mode and using a frame binning value of 2. Lowering this number of pre-averaged frames in AcquireMP requires a proportional increase in the number of averaged frames in post-processing in DiscoverMP to maintain default averaging time of ~100 ms.

Stage IV: Buffer background measurement

Timing 5 min

22) Fill the gasket with 10 µL buffer and focus using AUTOFOCUS. You can check that optimum focus has been found by manually translating the stage vertically up and down and verifying that the image sharpness is at a global maximum.

23) Check that the native image is clean (**Fig. 1b**) and free from bright spots which saturate the camera (**Fig. 4d**); if they are present, reposition the coverslip laterally and focus again; avoiding these spots will also result in lower image sharpness, which is in the range of 4-6% for a typical instrument and glass coverslip.

24) Check the ratiometric image for large buffer contaminants and fringes (**Fig. 4f, g**).

25) Record a movie for 60s. Avoid touching the instrument or inducing any vibrations during the measurement.

26) Plot mass histogram to confirm that the buffer does not contain impurities.

TROUBLESHOOTING

Stage V: Calibration

Timing 10 minutes

27) Immediately after a buffer background measurement, add 10 µl of 250x diluted calibrant Massference-p1 for a total 500x dilution to the same well, aspirating 3 times with the pipette.

28) Confirm that the glass-water interface is still in focus, if not refocus.

29) Record a 60s movie.

30) Process the movie of the calibrant with default thresholds and a total averaging time of 100 ms in DiscoverMP. Fit Gaussians to the first four peaks of Massference-p1 and assign masses of 86, 172, 258 and 344 kDa (**Fig. 1g**). Calibration slope should have R^2 > 0.99 (**Fig. 1h**).

31) Process the movie of the buffer background measurement in DiscoverMP and apply calibration.

- Check the histogram of the buffer movie, there will typically be <30 binding events total in a 1-minute movie of a clean buffer in the regular FoV for default thresholds and averaging time of 100 ms.
- After the calibration measurement, confirm there was no focus drift during the measurement. Focus drift could have resulted from vertical stage movement from the focus position while monitoring the sharpness to ensure that the sharpness decreases

in both directions. . If there is any evidence of focus drift, repeat the measurement and ensure stable mass reading over time (**Fig. 4h**).

Critical step: If accurate mass measurement is desired, confirm that buffer is clean and calibrate the mass to contrast conversion every two hours and at least at the start and end of the measurement series, (**Fig. 1g, h**).

TROUBLESHOOTING

Stage VI: Sample measurement

Timing 5 minutes/measurement

32) Use a coverslip with same surface chemistry as used for the calibration.
33) Prepare stock solutions in clean buffer (60 µM for BSA, 8.3 µM SARS-CoV-2 for antibody) and keep them on ice.
34) Dilute stock to working solutions at a concentration of 1.25 µM and keep it on ice.
35) Dilute working solution to 12.5 nM in 20 µL and incubate at room temperature to equilibrate the solution. Incubate for 2 minutes for BSA and antibodies. For quantifying the interactions, where possible consider the (likely) on and off rates of your analyte and set the incubation time accordingly. For quantification of interactions with slow association rates a longer incubation time is needed to assemble and equilibrate the complexes. Provided the complexes also have slow dissociation rates, manual rapid dilution into the gasket can be applied. In this scenario, the incubation concentration should be used for affinity quantification (see Jarmoskaite *et al.* for detailed guidance)[38].

Critical step. For reproducible measurements, the incubation time in the Eppendorf should be as similar as possible between repeats. Ensure that the sample is at room

temperature before adding to the gasket. Add buffer 2 minutes before measurement to minimize volume reduction due to evaporation. For working at concentrations <1 µM use PCR tubes instead of standard Eppendorf tubes to reduce protein adsorption to the tube walls. Aspirate quickly, to avoid protein deposition to the pipette tip surface, but prevent bubble formation.

36) During incubation period, move the microscope stage to a fresh gasket on the coverslip, add 4 µl of buffer and find focus.
37) Find an area of the coverslip with no saturated pixels in the native image, check the ratiometric movie and refocus on the spot intended for measurement.
38) Add 16 µl of protein to the 4 µl of buffer in the gasket, for a total concentration of 10 nM, aspirate 3 times to mix the solution. Avoid touching the coverslip surface or gasket with the pipette tip.

Critical step: For predominantly monomeric proteins such as BSA or SARS-CoV-2 antibodies, measure at monomeric concentrations of 2, 5, 10, 20 nM in the gasket to determine the optimal concentration. For oligomeric proteins or interacting protein mixtures, increase the monomeric concentration accordingly, to compensate the decrease in total particle counts due to the interactions. Measure 3 technical repeats at the previously determined optimal particle concentration.

39) Close the lid and start the acquisition immediately, recording for 60 s. Visually check if the ratiometric image is free from fringes (**Fig. 4f**), large buffer contaminants (**Fig. 4g**), and poor binding events (**Fig. 5a**).

TROUBLESHOOTING

Stage VII: Analysis of a single measurement

Timing 10 minutes/measurement

<Critical> It is always preferred to address any of the described issues at the acquisition phase, rather than during the data analysis.

40) Process movies using the same analysis settings as used for calibration and buffer background estimation.
41) Apply the mass calibration from the Massference-p1 measurement.
42) If the smallest expected peak (66 kDa peak of BSA) is not observed, increase the number of averaged frames in the measurement settings until it fully appears, i.e., does not increase in abundance if averaging time is further increased (**Fig. 7a**).
43) Check the presence of a low mass peak close to the detection limit that appears symmetrically in the positive and negative mass, indicating false positive detection (**Fig. 4a, b**). To address this, incrementally increase Threshold 1 until the number of events in the symmetric noise peaks is less than 10% of the number of events of each peak of interest (**Fig. 4c**).
44) Check that the spatial distribution of events is homogeneous. If regions with significantly higher density of events are detected, caused for example by wobblers, spatially mask those parts of the movie or/and use the nearest neighbor filter (**Fig. 5e, f**).
45) Check the temporal evolution of the standard deviation of the ratiometric frames. If spikes are detected, for example caused by movement of large impurities, exclude those parts of the movie (**Fig. 5d**).

46) Check if the peaks are baseline resolved in the mass histogram. If they are not, repeat measurement, or functionalize surface to improve binding.

47) Fit the peaks. The mass standard deviation of species ≤150 kDa should be ≤10 kDa. For BSA the expected masses of monomer and dimer are 66 kDa and 132 kDa, for the SARS-CoV-2 antibody the expected mass is 150 kDa, (**Fig. 6a, d** and **9b**).

Critical step: Always use the same analysis settings (thresholds and averaging time) for calibration and measurement. Optimal analysis settings depend on the type of measurement and can vary with the version of the analysis software.

TROUBLESHOOTING

Stage VIII: Analysis of multiple measurements or titration

Timing 10 min/measurement

48) For technical repeats, compare the counts in each peak, fluctuations in counts should be less than 25% (**Fig. 2**).

49) For a titration, analyse the measurements at different concentrations using the same analysis parameters, except for higher concentrations (for BSA > 10 nM), where the averaging time may be reduced to prevent overlapping PSFs.

50) Determine a suitable averaging time for accurate estimation of peak ratios (given by abundance of different species) in the selected concentration range. For BSA, plot the monomer/dimer ratio versus averaging time and then find the maximum of number of detected events and monomer:dimer ratio (**Fig. 7a, b**). Averaging time can be varied by changing the number of averaged frames in SETTINGS.

51) Plot the number of positive mass counts, monomer σ, and binding:unbinding ratio as a function of mass. The concentration range where counts scales linearly with

concentration should be identified. In this range, the error of the slope of a linear fit ≤ 5%, monomer σ ≤ 10 kDa (for species ≤ 150 kDa), and a binding:unbinding ratio ≤ 10 % should be achieved (**Fig. 6f,** and **Fig. 9c**).

52) Plot σ of the monomer and binding:unbinding ratio as a function of concentration. Both parameters should be constant at concentrations ≤ 10 nM (**Fig. 6 c, f**).

- **Critical step**: For quantification of protein-protein interactions with fast off-rates use the final concentration in the gasket in the calculations. For slow off-rates use the incubation concentration.

TROUBLESHOOTING

[H1] Troubleshooting

See **Table 1** and Figure 8 for troubleshooting.

[H1] Anticipated Results

Massference-p1 at 500x dilution from stock concentration exhibits peaks with sufficient events for accurate fitting at 86, 172, 258, and 344 kDa, with expected sigmas of 10, 13, 17, and 17 kDa, respectively (**Fig. 1g**). The calibration curve shows a mass to contrast ratio of -0.03 kDa^{-1}, typical for a well performing instrument, and the R^2 of the fit is 0.9997. Baseline resolution of BSA monomer and dimer is achieved at 10 nM on an APTES surface, with mass measurements of 66 and 132 kDa for the monomer and dimer, respectively (**Fig. 6a**). SARS-CoV-2-antibodies at 10 nM show substantial unbinding on glass and peak broadening due to inaccurate contrast estimation of poor binding events (**Fig. 9a**). High unbinding, which also leads to high local particle concentration above the surface thereby increasing the binding rate, can be reduced on APTES, which in turn provides access to higher measurement

concentrations (**Fig. 9b**). Such an increase of concentration range is important for measurement of lower affinity interactions[27]. The measured counts are linear in concentration up to 50 nM, confirming quantitative detection, and σ <10 kDa can be obtained at all these concentrations (**Fig 9c**).

Data availability

The datasets for all figures are available at 10.5281/zenodo.15800728. All data are shared under the CC BBY 4.0 license.

Table 1. Troubleshooting.

Stage	Step	Problem	Reason	Solution
Coverslip cleaning	3	Dirty coverslip surface after cleaning.	Glassware is dirty.	Clean glassware with glass detergent such as Hellmanex, rinse well with DI water.
			Due to the manufacturing process, both sides of the coverslip surface are not identical.	Compare native intensity of both orientations for two cleaned coverslips in the same batch and consistently use the lower intensity orientation.
Surface amination	14	Aggregates on APTES functionalized coverslips.	Contamination of glassware with APTES aggregates.	Always keep one specific beaker for APTES functionalisation, sonicate for 30 mins in acetone after functionalising.
			APTES stock solution contains aggregates, that developed over time due to contact with moisture from the air.	Use new APTES stock bottle. Alternatively, reduce incubation time with APTES to 1 min and do not sonicate after silanization, but instead twice wash by agitation for 1 min in a beaker with fresh acetone. Then blow dry the coverslips with nitrogen and crosslink APTES in the oven at 110°C for 2 hours. Then perform standard coverslip cleaning procedure.

Buffer background measurement	22	Gaskets leaking.	Gaskets have become dirty, compromising their liquid-tight adsorption onto the glass.	Use new gasket pack and cover the gaskets with film when stored to prevent dust.
				Clean gaskets by sonication in isopropanol and water followed by nitrogen drying.
	23,26	High numbers of events in buffer.	Buffer is contaminated.	Prepare fresh buffer. Alternatively, filter through a ≤ 0.22 µm filter, ensuring no additional impurities are introduced by the filter itself.
			Buffer contains macromolecules, such as PEG, detergent or glycerol.	Dilute buffer to reduce concentration of macromolecules, until there is no false positive detection in a buffer video. If this is not possible, and the analytes of interest have a high molecular weight, increase the strictness of the Welch's t-test ('Threshold 1' in DiscoverMP).
			False positive detection of background fluctuations.	Increase the strictness of the Welch's t-test ('Threshold 1' in DiscoverMP).
Calibration + Sample measurement	29,39	Focus is not stable during measurement.	Autofocussing feedback loop distorted by focus ring being discontinuous due to air bubbles	Lift the coverslip on one side and allow bubbles to move. If this does not resolve the issue, add more oil or clean objective and add oil again.

			between the objective and the coverslip.	
			Oil is on the stage.	Remove the coverslip, clean the stage with cotton buds and IPA. Change to new coverslip.
			Measurement buffer is not at room temperature.	Wait longer for the solution to equilibrate to room temperature before adding to the gasket.
			Measurement buffer not identical to buffer loaded for focusing, causing a difference in refractive index.	Use the same buffer for focusing and measurement. If same buffer cannot be used, manually fine refocus Z stage directly after sample addition, right before acquiring each movie.
			The volume in the gasket during focus finding was too low.	Find the minimum volume required for focusing by incremental addition of 1 µL of buffer until there is no more focus drift.
Sample measurement	39,40	No events detected.	Protein was lost due to deposition on the Eppendorf tube walls or pipette tips during storage, incubation, or dilution.	Use PCR tubes, mix quickly by aspirating up and down, reduce incubation times at low concentrations.
			Protein is aggregated.	Avoid making bubbles during dilutions avoid fast vortexing.

			Detection threshold is set too high.	Decrease the strictness of the Welch's t-test (lower the 'Threshold 1' in DiscoverMP).
			Analyte mass is below absolute detection limit (for TwoMP, 30 kDa).	Use analyte with mass higher than detection limit (for TwoMP >30 kDa).
			Back reflections from droplet create large signal (fringes).	Translate stage laterally until there are no fringes visible after adding the protein containing solution.
Analysis of a single measurement	41	Mass is inaccurate.	Mass calibration is inaccurate.	Apply mass calibration that is less than 2 hours old. Check that the mass to contrast slope is linear. Repeat measurement until $R^2 > 0.96$ is achieved.
			Species mass is close to detection limit, cutting off lower mass events. Measured mass will exceed nominal mass.	Decrease the strictness of the Welch's t-test (lower the 'Threshold 1' in DiscoverMP) and/or increase averaging time.
			Different buffer is used for measurement and calibration.	Use identical buffers for calibration and measurements.
	47	Mass peaks are too broad: σ >10 kDa for species	Focus drifts during measurement.	Exclude parts of the video where the standard deviation of the frame intensity drifts or redo measurement.

		≤150 kDa or >5% of the peak mass for larger species.	Event density is too high.	Reduce measurement concentrations until the total number of binding and unbinding events in the first minute does not exceed approximately 3.500 (small FoV), 6.000 (regular FoV) or 30.000 (large FoV).
			Surface binding is poor.	Reduce measurement concentration, functionalise batch of coverslips with fresh APTES, and/or change the buffer pH to change the protein charge and promote binding to the surface. If binding cannot be improved experimentally, use analysis tools such as nearest neighbour filtering, or spatial/temporal masking.
			Illumination is inhomogeneous.	Check if mass histograms resulting from different parts of the FoV in the same movie result in different average peak masses. If they differ by more than 2%, recalibrate the acquisition settings.
			Molecular heterogeneity.	Post-translational modification, such as glycosylation, can create a variety of species that are unresolvable, leading to peak broadening. This is an intrinsic property of the analyte.

Analysis of multiple measurements /titration	50	Inconsistent mass.	Incorrect focus position being used.	If using the automatic focus finder, check that the sharpness is at a global maximum by manually moving the stage up and down. Ensure coverslip is not touched by the pipette tip when adding the sample.
			Insufficient warmup time for the instrument to equilibrate.	Leave the instrument turned on with software running for a longer period (> 2 hours) prior to the measurement. Take regular calibration measurements during the day to account for contrast drift.
	48	Inconsistent counts.	Protein was lost due to adsorption on the Eppendorf tube walls during incubation or storage especially at low concentration.	Make a fresh dilution from the working solution (concentration >1 µM) or stock solution prior to each measurement.
			Inconsistent time between adding protein and starting acquisition, because the binding rate decays over time.	Time the delay between adding the sample and starting the acquisition and prioritize consistency over speed.
			Solution not fully mixed after adding	Rapidly aspirate several times in the gasket after adding the protein sample with pipette set to a volume

			protein sample to gasket.	> 50% of the total volume in the gasket. Make sure to not generate bubbles.

Author Contribution Statement

J.K, R.W, D.L, JC.T, J.B, K.I are co-first authors, they contributed equally. Everyone agrees their authorship orders can be exchanged on their CVs. All authors contributed to the conceptualisation. Data acquisition, analysis, interpretation, and writing of corresponding protocol section: Fig 1 JC.T; Fig 2, 3d, 4g-i D.L; Fig 3a-c,e-g J.B; Fig 4a-f R.W; Fig 5 K.I; Fig 6, 7, 9 JK; Fig 8 S.T. Writing of introduction S.T, review and editing of final manuscript S.T. and P.K. Supervision S.T and P.K.

Competing Interests Statement

P.K. is a nonexecutive director, shareholder of and consultant to Refeyn Ltd.

Acknowledgements

This work was funded by the European Research Council (ERC) Consolidator Grant PHOTOMASS 819593 (P.K, K.I), the Engineering and Physical Research Council (EPSRC) Leadership Fellowship EP/T03419X/1 (P.K, JC.T), the Biotechnology and Biological Sciences Research Council BB/W00349X/1, (P.K, S.T), the Wellcome Trust, Grant Number: 218514/Z/19/Z (R.W), UK Research and Innovation (UKRI) under the UK government's Horizon Europe funding guarantee through project Marie Skłodowska-Curie Actions (MSCA) Postdoctoral Fellowship NanoMassCreator (101062868) EP/X025713/1 (J.K), Clarendon scholarship, Menasseh Ben Israel scholarship, Kingsgate scholarship (D.L), EPRSC Doctoral Training Partnership (J.B).

For the purpose of Open Access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript (AAM) version arising from this submission.

The authors thank Konstantin Zouboulis, Francesca Naughton-Allen, Alexander O'Shea for feedback on the manuscript and discussion. The authors thank Manish Kushwah for providing the Dyn1 ΔPRD protein.

References

1. Young, G. *et al.* Quantitative mass imaging of single biological macromolecules. *Science* **360**, 423–427 (2018).
2. Asor, R., Loewenthal, D., Wee, R. van, Benesch, J. L. P. & Kukura, P. Mass Photometry. *Annu Rev Biophys* **54**, 379–399 (2025).
3. Asor, R. *et al.* Oligomerization-driven avidity correlates with SARS-CoV-2 cellular binding and inhibition. *Proc Natl Acad Sci U S A* **121**, e2403260121 (2024).
4. Olerinyova, A. *et al.* Mass Photometry of Membrane Proteins. *Chem* **7**, 224–236 (2021).
5. Sonn-Segev, A. *et al.* Quantifying the heterogeneity of macromolecular machines by mass photometry. *Nature Communications 2020 11:1* **11**, 1–10 (2020).
6. Fineberg, A., Surrey, T. & Kukura, P. Quantifying the Monomer–Dimer Equilibrium of Tubulin with Mass Photometry. *J Mol Biol* **432**, 6168–6172 (2020).
7. Zinder, J. C. *et al.* Shelterin is a dimeric complex with extensive structural heterogeneity. *Proc Natl Acad Sci U S A* **119**, e2201662119 (2022).
8. Kissling, V. M. *et al.* Mre11-Rad50 oligomerization promotes DNA double-strand break repair. *Nature Communications 2022 13:1* **13**, 1–16 (2022).
9. Liebthal, M., Kushwah, M. S., Kukura, P. & Dietz, K. J. Single molecule mass photometry reveals the dynamic oligomerization of human and plant peroxiredoxins. *iScience* **24**, (2021).
10. Hundt, N. *et al.* Direct observation of the molecular mechanism underlying protein polymerization. *Sci Adv* **8**, 31 (2022).
11. Wagner, C., Fuchsberger, F. F., Innthaler, B., Lemmerer, M. & Birner-Gruenberger, R. Quantification of Empty, Partially Filled and Full Adeno-Associated Virus Vectors Using Mass Photometry. *International Journal of Molecular Sciences 2023, Vol. 24, Page 11033* **24**, 11033 (2023).
12. Ebberink, E. H. T. M., Ruisinger, A., Nuebel, M., Thomann, M. & Heck, A. J. R. Assessing production variability in empty and filled adeno-associated viruses by single molecule mass analyses. *Mol Ther Methods Clin Dev* **27**, 491–501 (2022).
13. Häußermann, K., Young, G., Kukura, P. & Dietz, H. Dissecting FOXP2 Oligomerization and DNA Binding. *Angewandte Chemie International Edition* **58**, 7662–7667 (2019).
14. Bigelyte, G. *et al.* Miniature type V-F CRISPR-Cas nucleases enable targeted DNA modification in cells. *Nature Communications 2021 12:1* **12**, 1–8 (2021).
15. Acharya, A. *et al.* Distinct RPA domains promote recruitment and the helicase-nuclease activities of Dna2. *Nature Communications 2021 12:1* **12**, 1–15 (2021).

16. Hickman, A. B., Kailasan, S., Genzor, P., Haase, A. D. & Dyda, F. Casposase structure and the mechanistic link between DNA transposition and spacer acquisition by CRISPR-Cas. *Elife* **9**, (2020).
17. Gizardin-Fredon, H. *et al.* Denaturing mass photometry for rapid optimization of chemical protein-protein cross-linking reactions. *Nature Communications 2024 15:1* **15**, 1–13 (2024).
18. Wu, D. *et al.* Mass photometry: A powerful tool for carbohydrates-proteins conjugation monitoring and glycoconjugates molecular mass determination. *Glycoconj J* **40**, 401–412 (2023).
19. Young, J. W. *et al.* Characterization of membrane protein interactions by peptidisc-mediated mass photometry. *iScience* **27**, 108785 (2024).
20. Soltermann, F. *et al.* Quantifying Protein–Protein Interactions by Molecular Counting with Mass Photometry. *Angewandte Chemie International Edition* **59**, 10774–10779 (2020).
21. den Boer, M. A. *et al.* Comparative Analysis of Antibodies and Heavily Glycosylated Macromolecular Immune Complexes by Size-Exclusion Chromatography Multi-Angle Light Scattering, Native Charge Detection Mass Spectrometry, and Mass Photometry. *Anal Chem* **94**, 892–900 (2022).
22. Sendker, F. L. *et al.* Emergence of fractal geometries in the evolution of a metabolic enzyme. *Nature 2024 628:8009* **628**, 894–900 (2024).
23. Melo, L. *et al.* Size Distributions of Gold Nanoparticles in Solution Measured by Single-Particle Mass Photometry. *Journal of Physical Chemistry B* **125**, 12466–12475 (2021).
24. Lai, S. H., Tamara, S. & Heck, A. J. R. Single-particle mass analysis of intact ribosomes by mass photometry and Orbitrap-based charge detection mass spectrometry. *iScience* **24**, (2021).
25. 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. *Nature Methods 2021 18:10* **18**, 1247–1252 (2021).
26. Heermann, T., Steiert, F., Ramm, B., Hundt, N. & Schwille, P. Mass-sensitive particle tracking to elucidate the membrane-associated MinDE reaction cycle. *Nature Methods 2021 18:10* **18**, 1239–1246 (2021).
27. Kratochvíl, J., Asor, R., Helmi, S., Struwe, W. B. & Kukura, P. Lifting the Concentration Limit of Mass Photometry by PEG Nanopatterning. *Nano Lett* (2024).
28. Wu, D. & Piszczek, G. Standard protocol for mass photometry experiments. *Eur Biophys J* **50**, 403–409 (2021).
29. Kofinova, Z., Karunanithy, G., Ferreira, A. S. & Struwe, W. B. Measuring Protein-Protein Interactions and Quantifying Their Dissociation Constants with Mass Photometry. *Curr Protoc* **4**, e962 (2024).
30. Li, Y., Struwe, W. B. & Kukura, P. Single molecule mass photometry of nucleic acids. *Nucleic Acids Res* **48**, e97–e97 (2020).
31. Sülzle, J., Elfeky, L. & Manley, S. Surface passivation and functionalisation for mass photometry. *J Microsc* **295**, 14–20 (2024).
32. Ebberink, E. H. T. M. *et al.* Probing recombinant AAV capsid integrity and genome release after thermal stress by mass photometry. *Mol Ther Methods Clin Dev* **32**, 101293 (2024).
33. Van Wee, R. *et al.* NIPBL and STAG1 enable loop extrusion by providing differential DNA-cohesin affinity. *bioRxiv* 2024.10.22.619583 (2024).

34. Cole, D., Young, G., Weigel, A., Sebesta, A. & Kukura, P. Label-Free Single-Molecule Imaging with Numerical-Aperture-Shaped Interferometric Scattering Microscopy. *ACS Photonics* **4**, 211–216 (2017).
35. He, Y. *et al.* Multiscale modeling and analysis for high-fidelity interferometric scattering microscopy. *J Phys D Appl Phys* **54**, 274002 (2021).
36. Becker, J. *et al.* A Quantitative Description for Optical Mass Measurement of Single Biomolecules. *ACS Photonics* **10**, 2699–2710 (2023).
37. Tokuriki, N. *et al.* Protein folding by the effects of macromolecular crowding. *Protein Science* **13**, 125–133 (2004).
38. Jarmoskaite, I., Alsadhan, I., Vaidyanathan, P. P. & Herschlag, D. How to measure and evaluate binding affinities. *Elife* **9**, 1–34 (2020).
39. Claasen, M., Kofinova, Z., Contino, M. & Struwe, W. B. Analysis of Protein Complex Formation at Micromolar Concentrations by Coupling Microfluidics with Mass Photometry. *J Vis Exp* **2024**, (2024).

Figure 1. Principles of Mass Photometry. *(**a**) Measurement principle: Single proteins or complexes bind to a glass-water interface. The change in reflectivity is quantified to infer its mass. (**b**) The raw data is dominated by the static scattering from the glass roughness at the glass water interface. (**c**) Measurement of a diluted protein sample, Dyn1 ΔPRD. Intensity time trace of the landing event for a single 180 kDa species. The lower two panels are smoothed with a moving average. (**d**) Corresponding traces in the ratiometric data calculated with different averaging time windows. (**e**) Ratiometric frames at different time points (averaging time: 14 ms, scale bars: 1 µm). (**f**) Scatter plot of the extracted contrast and the landing time. (**g**) Contrast distribution for the calibrant Massference-p1. (**h**) The conversion between mass and contrast is determined based on the fits to the measured contrast and known masses of the calibration sample Massference-p1 (averaging time: 14 ms). (**i**) The mass distribution of the protein Dyn1 ΔPRD shows its oligomeric nature. The data in panels (a-h) were acquired using a small FoV (10.9 µm × 2.7 µm), while those in panel (i) used a regular FoV (10.9 µm x 4.3 µm). All data were measured using a total acquisition time of 1 minute.*

Figure 2. Measurement reproducibility. *Dataset of 8 consecutively measured replicates of Dyn1 ΔPRD at a monomer-equivalent concentration of 80 nM. (**a**) Detected events vs ratiometric contrast histograms, with a magnification of a single ratiometric contrast peak (inset). (**b**) Reproducibility metrics; consecutive measurements conducted according to our protocol can yield a total events count spread of about 30%, an oligomeric ratio spread of about 10% and a fitted mass slope spread of about 2%. (**c**) Normalized fitted ratiometric contrast values for the different oligomeric states; the fitted peak ratiometric contrast values for all oligomeric states vary by about 2% between replicates. (**d**) Log-scaled detected events vs ratiometric contrast histogram highlighting low-abundance, but still statistically significant species. All data were measured using a total acquisition time of 1 minute and large FoV (16.9 µm x 12.0 µm).*

Figure 3. Shot noise and how it affects quantitative detection and mass resolution*. (**a**) Effect of increased averaging time on ratiometric image (scalebars: 1 µm). (**b**) Detectability of Dyn1 ΔPRD monomer and dimer with increasing averaging time. The detection threshold was varied to obtain a*

*maximum number of counts while avoiding false positive events. (**c**) Quantification of binding events for a 150 kDa antibody (top) and a 42 kDa protein A (bottom). (**d**) Nominal vs measured mass determination for BSA (66 kDa), neutravidin (60 kDa), Protein A (42 kDa) and protein G (27 kDa) calibrated to Dyn1 ΔPRD. (**e**) Effect of increased averaging time on the Dyn1 ΔPRD histogram and tetramer FWHM. (**f**) Mass vs fitted peak standard deviation (sigma) for three repeats of Dyn1 ΔPRD using a large FoV (16.9 µm x 12.0 µm). (**g**) Effect of averaging time on Dyn1 ΔPRD tetramer sigma across standard FoVs (small 10.9 µm x 2.7 µm, regular 10.9 µm x 4.3 µm, large 16.9 µm x 12.0 µm). Data for other panels were collected using a small FoV (a-c, e), regular FoV (d) and large FoV (f). All data were measured using a total acquisition time of 1 minute.*

Figure 4. Noise contributions beyond shot noise. *(**a**) Quantification of the median standard deviation over space and time for increasing averaging times for a mass photometry video of PBS buffer. The difference between the datapoints (average ± standard deviation of 5 technical repeats) and the shot-noise limit as extrapolated from the first datapoint (solid line) constitutes excess noise (red). The noise floor (7.1 kDa) and corresponding averaging time (128 ms) are indicated. (**b**) Representative mass photometry images of the same FoV at three different averaging times. (**c**) MP histograms of a PBS measurement analyzed at 128 ms for various Welch's T-test filter thresholds ('Threshold 1' in DiscoverMP, default = 1.5) for particle detection. (**d**) Representative native MP FoV for a PBS buffer on a clean and dirty coverslip. The plot shows the intensity standard deviation across the horizontal dashed line for both (clean coverslip in blue, dirty coverslip in red). (**e**) Same as (**a**), but now also for a dirty coverslip (red diamonds) and PBS buffer containing 0.5% (w/v) PEG 8K (green squares). The inset shows a representative FoV of a ratiometric mass photometry video of PBS buffer containing 0.5% (w/v) poly-ethylene glycol (PEG) 8K at an averaging time of 128 ms. Contrast thresholds are as in (**b**). (**f**) Ratiometric mass photometry FoV of a buffer solution showing fringes created by back reflections from the droplet. (**g**) Ratiometric MP FoVs of a large sample impurity binding the coverslip and continuing to wobble at the same position. Snapshots are separated by 200 ms. Timetrace shows the mass standard deviation over the field, with the landing event indicated by the vertical dashed grey line. (**h**) Mass of detected events over time for a Dyn1 ΔPRD protein measurement. A linear fit reveals a trend of a 0.5-4 kDa (or 0.5% of the mass) decrease per minute in the mean mass value. (**i**) Local mean map of contrast values for a single mass photometry peak across the FoV, showing a nonuniformity of around 5%, with strong spatial correlation. (**j**) Influence of the reflectivity correction option in DiscoverMP on the standard deviation of Dyn1 ΔPRD protein peaks. The videos in panels (a-g) were acquired using a small FoV (10.9 µm × 2.7 µm), while those in panels (h-j) used a large FoV (16.9 µm × 12.0 µm). All data were measured using a total acquisition time of 1 minute. All scalebars: 1 µm.*

Figure 5. Accurate identification and quantification of binding events. *(**a**) Series of frames depicting distinct types of landing events that may occur over the course of a measurement. (**b**) Unprocessed*

*time series traces showing a binding and an unbinding event for a BSA dimer. (**c**) Corresponding unbinding event (red) ratiometrically processed over 80 ms averaging, highlighting the effect of poor binding on contrast estimation compared to the averaged trace from optimally binding events (black). Analytical methods can be employed to mitigate the effects of poor binding. (**d**) Selection of the temporal sections of the movie with low noise level (black), quantified using the framewise standard deviation; frame with highest noise is indicated by the arrow and shown in e. (**e**) Elevated noise in frame is caused by large aggregate (top, right). Frame shows wobbling particles, which were masked out by spatial filter (red squares). (**f**) Spatial distribution of events selected for analysis (black) and events excluded through temporal masking of frames with high standard deviation and spatial masking of frames with inhomogeneous landing density caused by wobbling particles (red). (**g**) Mass histogram of aged solution of Dyn1 ΔPRD exhibiting poor binding events (top). The corrected mass histogram by spatial and temporal masking (bottom); that was shown in panels d, f. All data were measured using an acquisition time of 2 minute and a regular FoV (10.9 µm × 4.3 µm). All scale bars: 1 µm.*

***Figure 6. Effect of concentration and surface chemistry on precision, accuracy, and quantification**. (**a, b**) Ratiometric frames and corresponding mass histograms from 1-minute measurements of BSA on glass at a concentration of 10 nM and 20 nM. (**c**). Dependence of binding counts (black, left y-axis), unbinding:binding ratio (red, middle y-axis) and monomer σ (blue, right y-axis) on BSA concentration on glass; the slope was estimated from datapoints in the 1-10 nM concentration range. (**d, e**) Frames and histograms of BSA on APTES at a concentration of 10 nM and 20 nM. (**f**) Binding counts (black, left y-axis), peak ratios (red, middle y-axis) and monomer σ (blue, right y-axis) as function of BSA concentration on APTES, slope was estimated in the 1-10 nM concentration range. (**g, h**) First and fourth minute of a measurement at 50 nM concentration and (**i**) events rates evolution during a 5-minute measurement. All data were acquired using small FoV (10.9 µm x 2.7 µm). All scale bars: 1 µm.*

***Figure 7. Effect of concentration and averaging window on quantifications of oligomer ratios.** (**a**, **b**) Number of detected binding events and monomer:dimer ratio for BSA as a function of averaging time at a concentration of 2 nM and 20 nM. (**c**) Ratiometric frames at 20 nM processed using different*

averaging windows of 6, 20, and 72 ms. All data were measured using a total acquisition time of 1 minute and small FoV (10.9 µm x 2.7 µm). Scale bars: 1 µm.

Figure 8. Flow diagram for the preparation, measurement and analysis of an MP experiment. *Troubleshooting suggestions are given for each step where relevant. This diagram is intended as near-exhaustive description of potential issues that could arise in an MP measurement. RT is room temperature, (N) is native and (R) ratiometric image.*

Figure 9. Titration measurement of SARS-CoV-2 antibodies. *Ratiometric frames and corresponding mass histograms of measurements at a concentration of 10 nM on* ***(a)*** *glass and* ***(b)*** *APTES.* ***(c)*** *Concentration dependence of counts, unbinding/binding ratio and monomer σ, which was measured on APTES surface; the slope was estimated in the concentration range of 1-10 nM. All data were measured using a total acquisition time of 1 minute and small FoV (10.9 µm x 4.3 µm). Scale bars: 1 µm.*